

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
29 March 2001 (29.03.2001)

PCT

(10) International Publication Number
WO 01/21189 A1(51) International Patent Classification: **A61K 38/00**,
38/04, 38/08, 38/10, 39/29, 39/295

(21) International Application Number: PCT/US00/19774

(22) International Filing Date: 19 July 2000 (19.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/357,737 19 July 1999 (19.07.1999) US(71) Applicant (for all designated States except US): **EPIMUNE INC.** [US/US]; 5820 Nancy Ridge Drive, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SETTE, Alessandro** [IT/US]; 5551 Linda Rosa Avenue, La Jolla, CA 92037 (US); **SIDNEY, John** [US/US]; 4218 Corte de la Siena, San Diego, CA 92130 (US); **SOUTHWOOD, Scott** [US/US]; 10679 Strathmore Drive, Santee, CA 92071 (US); **LIVINGSTON, Brian, D.** [US/US]; 13555 Chaco Court, San Diego, CA 92129 (US); **CHESNUT, Robert** [US/US]; 1473 Kings Cross Drive, Cardiff-by-the-Sea, CA 92007 (US); **BAKER, Denise, Marie** [US/US]; 11575 Caminito LaBar #21, San Diego, CA 92126 (US); **CELIS, Esteban** [US/US]; 3683 Wright Road S.W., Rochester,MN 55902 (US). **KUBO, Ralph, T.** [US/US]; 6921 Pear Tree Drive, Carlsbad, CA 92009 (US); **GREY, Howard, M.** [US/US]; 1461 Caminito Batea, La Jolla, CA 92037 (US).(74) Agents: **LOCKYER, Jean, M.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

(81) Designated States (national): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GI, GM, KR, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— With international search report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS C VIRUS USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

(57) Abstract: This invention uses our knowledge of the mechanisms by which antigen is recognized by T cells to identify and prepare HCV epitopes, and to develop epitope-based vaccines directed towards HCV. More specifically, this application communicates our discovery of pharmaceutical compositions and methods of use in the prevention and treatment of HCV infection.

WO 01/21189 A1

**INDUCING CELLULAR IMMUNE RESPONSES TO HEPATITIS C VIRUS
USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS**

5

FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

- 10 This invention was funded, in part, by the United States government under grants with the National Institutes of Health. The U.S. government has certain rights in this invention.

INDEX

- 15 I. Background of the Invention
 II. Summary of the Invention
 III. Brief Description of the Figures
 IV. Detailed Description of the Invention
 A. Definitions
20 B. Stimulation of CTL and HTL responses
 C. Binding Affinity of Peptide Epitopes for HLA Molecules
 D. Peptide Epitope Binding Motifs and Supermotifs
 1. HLA-A1 supermotif
 2. HLA-A2 supermotif
25 3. HLA-A3 supermotif
 4. HLA-A24 supermotif
 5. HLA-B7 supermotif
 6. HLA-B27 supermotif
 7. HLA-B44 supermotif
30 8. HLA-B58 supermotif
 9. HLA-B62 supermotif
 10. HLA-A1 motif

- 11. HLA-A2.1 motif
- 12. HLA-A3 motif
- 13. HLA-A11 motif
- 14. HLA-A24 motif
- 5 15. HLA-DR-1-4-7 supermotif
- 16. HLA-DR3 motifs
- E. Enhancing Population Coverage of the Vaccine
- F. Immune Response-Stimulating Peptide Epitope Analogs
- G. Computer Screening of Protein Sequences from Disease-Related Antigens
- 10 for Supermotif- or Motif-Containing Epitopes
- H. Preparation of Peptide Epitopes
- I. Assays to Detect T-Cell Responses
- J. Use of Peptide Epitopes for Evaluating Immune Responses
- K. Vaccine Compositions
- 15 1. Minigene Vaccines
- 2. Combinations of CTL Peptides with Helper Peptides
- L. Administration of Vaccines for Therapeutic or Prophylactic Purposes
- M. Kits
- V. Examples
- 20 VI. Claims
- VII. Abstract

1. BACKGROUND OF THE INVENTION

- Hepatitis C virus (HCV) infection is a global human health problem with
- 25 approximately 150,000 new reported cases each year in the U.S. alone. HCV is a single stranded RNA virus, and is the etiologic agent identified in most cases of non-A, non-B post-transfusion and post-transplant hepatitis, and is a common cause of acute sporadic hepatitis (Choo *et al.*, *Science* 244:359, 1989; Kuo *et al.*, *Science* 244:362, 1989; and Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989). It is estimated that more
- 30 than 50% of patients infected with HCV become chronically infected and, of those, 20% develop cirrhosis of the liver within 20 years (Davis *et al.*, *New Engl. J. Med.* 321:1501,

1989; Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989; Alter *et al.*, *New Engl. J. Med.* 327:1899, 1992; and Dienstag, J. L. *Gastroenterology* 85:430, 1983).

Moreover, the only therapy available for treatment of HCV infection is interferon- α .

Most patients are unresponsive, however, and among the responders, there is a high

- 5 recurrence rate within 6-12 months of cessation of treatment (Liang *et al.*, *J. Med. Virol.* 40:69, 1993). Ribavirin, a guanosine analog with a broad spectrum activity against many RNA and DNA viruses, has been shown in clinical trials to be effective against chronic HCV infection when used in combination with interferon- α (*see, e.g.*, Poynard *et al.*, *Lancet* 352:1426-1432, 1998; Reichard *et al.*, *Lancet* 351:83-87, 1998) However, the
10 response rate is still well below 50%.

- Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections *in vivo* (Oldstone *et al.*, *Nature* 321:239, 1989; Jamieson *et al.*, *J. Virol.* 61:3930, 1987; Yap *et al.*, *Nature* 273:238, 1978; Lukacher *et al.*, *J. Exp. Med.* 160:814,
15 1994; McMichael *et al.*, *N. Engl. J. Med.* 309:13, 1983; Sethi *et al.*, *J. Gen. Virol.* 64:443, 1983; Watari *et al.*, *J. Exp. Med.* 165:459, 1987; Yasukawa *et al.*, *J. Immunol.* 143:2051, 1989; Tigges *et al.*, *J. Virol.* 66:1622, 1993; Reddenhase *et al.*, *J. Virol.* 55:263, 1985; Quinnan *et al.*, *N. Engl. J. Med.* 307:6, 1982). HLA class I molecules are expressed on the surface of almost all nucleated cells. Following intracellular processing of antigens,
20 epitopes from the antigens are presented as a complex with the HLA class I molecules on the surface of such cells. CTL recognize the peptide-HLA class I complex, which then results in the destruction of the cell bearing the HLA-peptide complex directly by the CTL and/or via the activation of non-destructive mechanisms *e.g.*, the production of interferon, that inhibit viral replication.

- 25 In view of the heterogeneous immune response observed with HCV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple HCV epitopes appears to be important for the development of an efficacious vaccine against HCV. There is a need, however, to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HCV
30 infection.

The information provided in this section is intended to disclose the presently understood state of the art as of the filing date of the present application. Information is included in this section which was generated subsequent to the priority date of this

application. Accordingly, information in this section is not intended, in any way, to delineate the priority date for the invention.

II. SUMMARY OF THE INVENTION

5 This invention applies our knowledge of the mechanisms by which antigen is recognized by T cells, for example, to develop epitope-based vaccines directed towards HCV. More specifically, this application communicates our discovery of specific epitope pharmaceutical compositions and methods of use in the prevention and treatment of HCV infection.

10 Upon development of appropriate technology, the use of epitope-based vaccines has several advantages over current vaccines, particularly when compared to the use of whole antigens in vaccine compositions. There is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. The epitopes for inclusion in an epitope-based vaccine
15 are selected from conserved regions of viral or tumor-associated antigens, which thereby reduces the likelihood of escape mutants. Furthermore, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines.

An additional advantage of an epitope-based vaccine approach is the ability to combine selected epitopes (CTL and HTL), and further, to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune
20 response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole
25 protein antigens, which might have their own intrinsic biological activity, is eliminated.

An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen. Thus, patient-by-patient variability in the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from that pathogen in a vaccine composition. A
30 "pathogen" may be an infectious agent or a tumor associated molecule.

One of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics, however, has been the extreme polymorphism of HLA molecules. To date, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used

that are specific for HLA molecules corresponding to each individual HLA allele, therefore, impractically large numbers of epitopes would have to be used in order to cover ethnically diverse populations. Thus, there has existed a need for peptide epitopes that are bound by multiple HLA antigen molecules for use in epitope-based vaccines. The greater the number of HLA antigen molecules bound, the greater the breadth of population coverage by the vaccine.

Furthermore, as described herein in greater detail, a need has existed to modulate peptide binding properties, for example, so that peptides that are able to bind to multiple HLA antigens do so with an affinity that will stimulate an immune response.

- 10 Identification of epitopes restricted by more than one HLA allele at an affinity that correlates with immunogenicity is important to provide thorough population coverage, and to allow the elicitation of responses of sufficient vigor to prevent or clear an infection in a diverse segment of the population. Such a response can also target a broad array of epitopes. The technology disclosed herein provides for such favored immune responses.

- 15 In a preferred embodiment, epitopes for inclusion in vaccine compositions of the invention are selected by a process whereby protein sequences of known antigens are evaluated for the presence of motif or supermotif-bearing epitopes. Peptides corresponding to a motif- or supermotif-bearing epitope are then synthesized and tested for the ability to bind to the HLA molecule that recognizes the selected motif. Those peptides that bind at an intermediate or high affinity *i.e.*, an IC_{50} (or a K_D value) of 500 nM or less for HLA class I molecules or 1000 nM or less for HLA class II molecules, are further evaluated for their ability to induce a CTL or HTL response. Immunogenic peptide epitopes are selected for inclusion in vaccine compositions.

- 20 Supermotif-bearing peptides may additionally be tested for the ability to bind to multiple alleles within the HLA supertype family. Moreover, peptide epitopes may be analogued to modify binding affinity and/or the ability to bind to multiple alleles within an HLA supertype.

- 25 The invention also includes an embodiment comprising a method for monitoring or evaluating an immune response to HCV in a patient having a known HLA-type, the method comprising incubating a T lymphocyte sample from the patient with a peptide composition comprising an HCV epitope consisting essentially of an amino acid sequence described in Tables VII to Table XX or Table XXII which binds the product of at least one HLA allele present in said patient, and detecting for the presence of a T lymphocyte

that binds to the peptide. A CTL peptide epitope may, for example, comprise a tetrameric complex.

An alternative modality for defining the peptide epitopes in accordance with the invention is to recite the physical properties, such as length; primary structure; or charge, which are correlated with binding to a particular allele-specific HLA molecule or group of allele-specific HLA molecules. A further modality for defining peptide epitopes is to recite the physical properties of an HLA binding pocket, or properties shared by several allele-specific HLA binding pockets (e.g. pocket configuration and charge distribution) and reciting that the peptide epitope fits and binds to said pocket or pockets.

As will be apparent from the discussion below, other methods and embodiments are also contemplated. Further, novel synthetic peptides produced by any of the methods described herein are also part of the invention.

III. BRIEF DESCRIPTION OF THE FIGURES

Figure 1: Figure 1 provides a graph of total frequency of genotypes as a function of the number of HCV candidate epitopes bound by HLA-A and B molecules, in an average population.

Figure 2: Figure 2 illustrates the position of peptide epitopes in an experimental model minigene construct.

IV. DETAILED DESCRIPTION OF THE INVENTION

The peptide epitopes and corresponding nucleic acid compositions of the present invention are useful for stimulating an immune response to HCV by stimulating the production of CTL or HTL responses. The peptide epitopes, which are derived directly or indirectly from native HCV amino acid sequences, are able to bind to HLA molecules and stimulate an immune response to HCV. The complete polypeptide sequence from HCV and its variants can be obtained from Genbank. Peptide epitopes and analogs thereof can also be readily determined from sequence information that may subsequently be discovered for heretofore unknown variants of HCV, as will be clear from the disclosure provided below.

The peptide epitopes of the invention have been identified in a number of ways, as will be discussed below. Also discussed in greater detail is that analog peptides have been derived and the binding activity for HLA molecules modulated by modifying specific amino acid residues to create peptide analogs exhibiting altered immunogenicity.

Further, the present invention provides compositions and combinations of compositions that enable epitope-based vaccines that are capable of interacting with HLA molecules encoded by various genetic alleles to provide broader population coverage than prior vaccines.

5

IV.A. Definitions

The invention can be better understood with reference to the following definitions, which are listed alphabetically:

10 A "computer" or "computer system" generally includes: a processor; at least one information storage/retrieval apparatus such as, for example, a hard drive, a disk drive or a tape drive; at least one input apparatus such as, for example, a keyboard, a mouse, a touch screen, or a microphone; and display structure. Additionally, the computer may include a communication channel in communication with a network. Such a computer may include more or less than what is listed above.

15 "Cross-reactive binding" indicates that a peptide is bound by more than one HLA molecule; a synonym is degenerate binding.

A "cryptic epitope" elicits a response by immunization with an isolated peptide, but the response is not cross-reactive *in vitro* when intact whole protein which comprises the epitope is used as an antigen.

20 A "dominant epitope" is an epitope that induces an immune response upon immunization with a whole native antigen (see, e.g., Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993). Such a response is cross-reactive *in vitro* with an isolated peptide epitope.

25 With regard to a particular amino acid sequence, an "epitope" is a set of amino acid residues which is involved in recognition by a particular immunoglobulin, or in the context of T cells, those residues necessary for recognition by T cell receptor proteins and/or Major Histocompatibility Complex (MHC) receptors. In an immune system setting, *in vivo* or *in vitro*, an epitope is the collective features of a molecule, such as primary, secondary and tertiary peptide structure, and charge, that together form a site
30 recognized by an immunoglobulin, T cell receptor or HLA molecule. Throughout this disclosure epitope and peptide are often used interchangeably.

It is to be appreciated that protein or peptide molecules that comprise an epitope of the invention as well as additional amino acid(s) are still within the bounds of the invention. In certain embodiments, there is a limitation on the length of a peptide of the

invention which is not otherwise a construct. An embodiment that is length-limited occurs when the protein/peptide comprising an epitope of the invention comprises a region (i.e., a contiguous series of amino acids) having 100% identity with a native sequence. In order to avoid the definition of epitope from reading, e.g., on whole natural molecules, there is a limitation on the length of any region that has 100% identity with a native peptide sequence. Thus, for a peptide comprising an epitope of the invention and a region with 100% identity with a native peptide sequence (and is not otherwise a construct), the region with 100% identity to a native sequence generally has a length of: less than or equal to 600 amino acids, often less than or equal to 500 amino acids, often less than or equal to 400 amino acids, often less than or equal to 250 amino acids, often less than or equal to 100 amino acids, often less than or equal to 85 amino acids, often less than or equal to 75 amino acids, often less than or equal to 65 amino acids, and often less than or equal to 50 amino acids. In certain embodiments, an "epitope" of the invention is comprised by a peptide having a region with less than 51 amino acids that has 100% identity to a native peptide sequence, in any increment of (49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5) down to 5 amino acids.

Accordingly, peptide or protein sequences longer than 600 amino acids are within the scope of the invention, so long as they do not comprise any contiguous sequence of more than 600 amino acids that have 100% identity with a native peptide sequence, if they are not otherwise a construct. For any peptide that has five contiguous residues or less that correspond to a native sequence, there is no limitation on the maximal length of that peptide in order to fall within the scope of the invention. It is presently preferred that a CTL epitope be less than 600 residues long in any increment down to eight amino acid residues.

"Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (*see, e.g., Stites, et al., IMMUNOLOGY*, 8TH ED., Lange Publishing, Los Altos, CA (1994)).

An "HLA supertype or family", as used herein, describes sets of HLA molecules grouped on the basis of shared peptide-binding specificities. HLA class I molecules that share somewhat similar binding affinity for peptides bearing certain amino acid motifs are grouped into HLA superotypes. The terms HLA superfamily, HLA supertype family, HLA family, and HLA xx-like supertype molecules (where xx denotes a particular HLA type), are synonyms.

Throughout this disclosure, results are expressed in terms of "IC₅₀'s." IC₅₀ is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (i.e., limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. Assays for determining binding are described in detail, e.g., in PCT publications WO 94/20127 and WO 94/03205. It should be noted that IC₅₀ values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents used (e.g., HLA preparation, etc.). For example, excessive concentrations of HLA molecules will increase the apparent measured IC₅₀ of a given ligand.

Alternatively, binding is expressed relative to a reference peptide. Although as a particular assay becomes more, or less, sensitive, the IC₅₀'s of the peptides tested may change somewhat, the binding relative to the reference peptide will not significantly change. For example, in an assay run under conditions such that the IC₅₀ of the reference peptide increases 10-fold, the IC₅₀ values of the test peptides will also shift approximately 10-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good, intermediate, weak, or negative binder is generally based on its IC₅₀, relative to the IC₅₀ of a standard peptide.

Binding may also be determined using other assay systems including those using: live cells (e.g., Ceppellini *et al.*, *Nature* 339:392, 1989; Christnick *et al.*, *Nature* 352:67, 1991; Busch *et al.*, *Int. Immunol.* 2:443, 1990; Hill *et al.*, *J. Immunol.* 147:189, 1991; del Guercio *et al.*, *J. Immunol.* 154:685, 1995), cell free systems using detergent lysates (e.g., Cerundolo *et al.*, *J. Immunol.* 21:2069, 1991), immobilized purified MHC (e.g., Hill *et al.*, *J. Immunol.* 152, 2890, 1994; Marshall *et al.*, *J. Immunol.* 152:4946, 1994), ELISA systems (e.g., Reay *et al.*, *EMBO J.* 11:2829, 1992), surface plasmon resonance (e.g., Khilko *et al.*, *J. Biol. Chem.* 268:15425, 1993); high flux soluble phase assays (Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994), and measurement of class I MHC stabilization or assembly (e.g., Ljunggren *et al.*, *Nature* 346:476, 1990; Schumacher *et al.*, *Cell* 62:563, 1990; Townsend *et al.*, *Cell* 62:285, 1990; Parker *et al.*, *J. Immunol.* 149:1896, 1992).

As used herein, "high affinity" with respect to HLA class I molecules is defined as binding with an IC₅₀, or K_D value, of 50 nM or less; "intermediate affinity" is binding with an IC₅₀ or K_D value of between about 50 and about 500 nM. "High affinity" with respect to binding to HLA class II molecules is defined as binding with an IC₅₀ or K_D value of 100 nM or less; "intermediate affinity" is binding with an IC₅₀ or K_D value of between about 100 and about 1000 nM.

The terms "identical" or percent "identity," in the context of two or more peptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using a sequence comparison algorithm or by manual alignment and visual inspection.

An "immunogenic peptide" or "peptide epitope" is a peptide that comprises an allele-specific motif or supermotif such that the peptide will bind an HLA molecule and induce a CTL and/or HTL response. Thus, immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and thereafter inducing an HLA-restricted cytotoxic or helper T cell response to the antigen from which the immunogenic peptide is derived.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated with the peptides in their *in situ* environment. An "isolated" epitope refers to an epitope that does not include the whole sequence of the antigen or polypeptide from which the epitope was derived. Typically the "isolated" epitope does not have attached thereto additional amino acids that result in a sequence that has 100% identity with a native sequence. The native sequence can be a sequence such as a tumor-associated antigen from which the epitope is derived.

"Major Histocompatibility Complex" or "MHC" is a cluster of genes that plays a role in control of the cellular interactions responsible for physiologic immune responses. In humans, the MHC complex is also known as the HLA complex. For a detailed description of the MHC and HLA complexes, see, Paul, FUNDAMENTAL IMMUNOLOGY, 3RD ED., Raven Press, New York, 1993.

The term "motif" refers to the pattern of residues in a peptide of defined length, usually a peptide of from about 8 to about 13 amino acids for a class I HLA motif and from about 6 to about 25 amino acids for a class II HLA motif, which is recognized by a particular HLA molecule. Peptide motifs are typically different for each protein encoded by each human HLA allele and differ in the pattern of the primary and secondary anchor residues.

A "negative binding residue" is an amino acid which, if present at certain positions (typically not primary anchor positions) in a peptide epitope, results in decreased binding affinity of the peptide for the peptide's corresponding HLA molecule.

A "non-native" sequence or "construct" refers to a sequence that is not found in in nature ("non-naturally occurring"). Such sequences include, *e.g.*, peptides that are lipidated or otherwise modified and polypeptidic compositions that contain epitopes that are non contiguous in a native protein sequence.

- 5 The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other, typically by peptide bonds between the α -amino and carboxyl groups of adjacent amino acids. The preferred CTL-inducing peptides of the invention are 13 residues or less in length and usually consist of between about 8 and about 11 residues, preferably 9 or 10 residues. The preferred HTL-inducing oligopeptides are less than about 50 residues in length and usually consist of between about 6 and about 30 residues, more usually between about 12 and 25, and often between about 15 and 20 residues.

"Pharmaceutically acceptable" refers to a generally non-toxic, inert, and/or physiologically compatible composition.

- 15 A "pharmaceutical excipient" comprises a material such as an adjuvant, a carrier, pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservative, and the like.

- A "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One to three, usually two, primary anchor residues within a peptide of defined length generally defines a "motif" for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding grooves of an HLA molecule, with their side chains buried in specific pockets of the binding grooves themselves. In one embodiment, the primary anchor residues are located at position 2 (from the amino terminal position) and at the carboxyl terminal position of a 9-residue peptide epitope in accordance with the invention. The primary anchor positions for each motif and supermotif are set forth in Table 1. For example, analog peptides can be created by altering the presence or absence of particular residues in these primary anchor positions. Such analogs are used to modulate the binding affinity of a peptide comprising a particular motif or supermotif.
- 20
25
30

"Promiscuous recognition" is where a distinct peptide is recognized by the same T cell clone in the context of various HLA molecules. Promiscuous recognition or binding is synonymous with cross-reactive binding.

A "protective immune response" or "therapeutic immune response" refers to a CTL and/or an HTL response to an antigen derived from an infectious agent or a tumor antigen, which prevents or at least partially arrests disease symptoms or progression. The immune response may also include an antibody response which has been facilitated by the stimulation of helper T cells.

The term "residue" refers to an amino acid or amino acid mimetic incorporated into an oligopeptide by an amide bond or amide bond mimetic.

A "secondary anchor residue" is an amino acid at a position other than a primary anchor position in a peptide which may influence peptide binding. A secondary anchor residue occurs at a significantly higher frequency amongst bound peptides than would be expected by random distribution of amino acids at one position. The secondary anchor residues are said to occur at "secondary anchor positions." A secondary anchor residue can be identified as a residue which is present at a higher frequency among high or intermediate affinity binding peptides, or a residue otherwise associated with high or intermediate affinity binding. For example, analog peptides can be created by altering the presence or absence of particular residues in these secondary anchor positions. Such analogs are used to finely modulate the binding affinity of a peptide comprising a particular motif or supermotif.

A "subdominant epitope" is an epitope which evokes little or no response upon immunization with whole antigens which comprise the epitope, but for which a response can be obtained by immunization with an isolated peptide, and this response (unlike the case of cryptic epitopes) is detected when whole protein is used to recall the response *in vitro* or *in vivo*.

A "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles. Preferably, a supermotif-bearing peptide is recognized with high or intermediate affinity (as defined herein) by two or more HLA antigens.

"Synthetic peptide" refers to a peptide that is man-made using such methods as chemical synthesis or recombinant DNA technology.

As used herein, a "vaccine" is a composition that contains one or more peptides of the invention. There are numerous embodiments of vaccines in accordance with the invention, such as by a cocktail of one or more peptides; one or more epitopes of the invention comprised by a polyepitopic peptide; or nucleic acids that encode such peptides or polypeptides, e.g., a minigene that encodes a polyepitopic peptide. The "one or more peptides" can include, e.g., at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,

- 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 or 100 or more peptides of the invention. The peptides or polypeptides can optionally be modified, such as by lipidation, addition of targeting or other sequences. HLA class I-binding peptides of the invention can be admixed with, or
- 5 linked to, HLA class II-binding peptides, to facilitate activation of both cytotoxic T lymphocytes and helper T lymphocytes. Vaccines can also comprise peptide-pulsed antigen presenting cells, e.g., dendritic cells.

- The nomenclature used to describe peptide compounds follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the
- 10 carboxyl group to the right (the C-terminus) of each amino acid residue. When amino acid residue positions are referred to in a peptide epitope they are numbered in an amino to carboxyl direction with position one being the position closest to the amino terminal end of the epitope, or the peptide or protein of which it may be a part. In the formulae representing selected specific embodiments of the present invention, the amino- and
- 15 carboxyl-terminal groups, although not specifically shown, are in the form they would assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or single letter designations. The L-form of an amino acid residue is represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids
- 20 having D-forms is represented by a lower case single letter or a lower case three letter symbol. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or G. Symbols for the amino acids are shown below.

Single Letter Symbol	Three Letter Symbol	Amino Acids
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

IV.B. Stimulation of CTL and HTL responses

The mechanism by which T cells recognize antigens has been delineated during the past ten years. Based on our understanding of the immune system we have developed efficacious peptide epitope vaccine compositions that can induce a therapeutic or prophylactic immune response to HCV in a broad population. For an understanding of the value and efficacy of the claimed compositions, a brief review of immunology-related technology is provided.

A complex of an HLA molecule and a peptidic antigen acts as the ligand recognized by HLA-restricted T cells (Buus, S. *et al.*, *Cell* 47:1071, 1986; Babbitt, B. P. *et al.*, *Nature* 317:359, 1985; Townsend, A. and Bodmer, H., *Annu. Rev. Immunol.* 7:601,

- 1989; Germain, R. N., *Annu. Rev. Immunol.* 11:403, 1993). Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues that correspond to motifs required for specific binding to HLA antigen molecules have been identified and are described herein and are set forth in Tables I, II, and III (see also, e.g., Southwood, *et al.*, *J. Immunol.* 160:3363, 1998; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995; Rammensee *et al.*, SYFPEITHI, access via web at : <http://134.2.96.221/scripts.hlaserver.dll/home.htm>; Sette, A. and Sidney, J. *Curr. Opin. Immunol.* 10:478, 1998; Engelhard, V. H., *Curr. Opin. Immunol.* 6:13, 1994; Sette, A. and Grey, H. M., *Curr. Opin. Immunol.* 4:79, 1992;
- 10 Sinigaglia, F. and Hammer, J. *Curr. Biol.* 6:52, 1994; Ruppert *et al.*, *Cell* 74:929-937, 1993; Kondo *et al.*, *J. Immunol.* 155:4307-4312, 1995; Sidney *et al.*, *J. Immunol.* 157:3480-3490, 1996; Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; Sette, A. and Sidney, J. *Immunogenetics*, in press, 1999).

- Furthermore, x-ray crystallographic analysis of HLA-peptide complexes has revealed pockets within the peptide binding cleft of HLA molecules which accommodate, in an allele-specific mode, residues borne by peptide ligands; these residues in turn determine the HLA binding capacity of the peptides in which they are present. (See, e.g., Madden, D.R. *Annu. Rev. Immunol.* 13:587, 1995; Smith, *et al.*, *Immunity* 4:203, 1996; Fremont *et al.*, *Immunity* 8:305, 1998; Stern *et al.*, *Structure* 2:245, 1994; Jones, E.Y. *Curr. Opin. Immunol.* 9:75, 1997; Brown, J. H. *et al.*, *Nature* 364:33, 1993; Guo, H. C. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:8053, 1993; Guo, H. C. *et al.*, *Nature* 360:364, 1992; Silver, M. L. *et al.*, *Nature* 360:367, 1992; Matsumura, M. *et al.*, *Science* 257:927, 1992; Madden *et al.*, *Cell* 70:1035, 1992; Fremont, D. H. *et al.*, *Science* 257:919, 1992; Saper, M. A., Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991.)
- 20

- 25 Accordingly, the definition of class I and class II allele-specific HLA binding motifs, or class I or class II supermotifs allows identification of regions within a protein that have the potential of binding particular HLA antigen(s).

- The present inventors have found that the correlation of binding affinity with immunogenicity, which is disclosed herein, is an important factor to be considered when evaluating candidate peptides. Thus, by a combination of motif searches and HLA-peptide binding assays, candidates for epitope-based vaccines have been identified. After determining their binding affinity, additional confirmatory work can be performed to select, amongst these vaccine candidates, epitopes with preferred characteristics in terms of population coverage, antigenicity, and immunogenicity.
- 30

Various strategies can be utilized to evaluate immunogenicity, including:

- 1) Evaluation of primary T cell cultures from normal individuals (*see, e.g.,* Wentworth, P. A. *et al.*, *Mol. Immunol.* 32:603, 1995; Celis, E. *et al.*, *Proc. Natl. Acad. Sci. USA* 91:2105, 1994; Tsai, V. *et al.*, *J. Immunol.* 158:1796, 1997; Kawashima, I. *et al.*, *Human Immunol.* 59:1, 1998); This procedure involves the stimulation of peripheral blood lymphocytes (PBL) from normal subjects with a test peptide in the presence of antigen presenting cells *in vitro* over a period of several weeks. T cells specific for the peptide become activated during this time and are detected using, *e.g.*, a ^{51}Cr -release assay involving peptide sensitized target cells.
- 2) Immunization of HLA transgenic mice (*see, e.g.,* Wentworth, P. A. *et al.*, *J. Immunol.* 26:97, 1996; Wentworth, P. A. *et al.*, *Int. Immunol.* 8:651, 1996; Alexander, J. *et al.*, *J. Immunol.* 159:4753, 1997); In this method, peptides in incomplete Freund's adjuvant are administered subcutaneously to HLA transgenic mice. Several weeks following immunization, splenocytes are removed and cultured *in vitro* in the presence of test peptide for approximately one week. Peptide-specific T cells are detected using, *e.g.*, a ^{51}Cr -release assay involving peptide sensitized target cells and target cells expressing endogenously generated antigen.
- 3) Demonstration of recall T cell responses from immune individuals who have effectively been vaccinated, recovered from infection, and/or from chronically infected patients (*see, e.g.,* Rehmann, B. *et al.*, *J. Exp. Med.* 181:1047, 1995; Doolan, D. L. *et al.*, *Immunity* 7:97, 1997; Bertoni, R. *et al.*, *J. Clin. Invest.* 100:503, 1997; Threlkeld, S. C. *et al.*, *J. Immunol.* 159:1648, 1997; Diepolder, H. M. *et al.*, *J. Virol.* 71:6011, 1997). In applying this strategy, recall responses are detected by culturing PBL from subjects that have been naturally exposed to the antigen, for instance through infection, and thus have generated an immune response "naturally", or from patients who were vaccinated against the infection. PBL from subjects are cultured *in vitro* for 1-2 weeks in the presence of test peptide plus antigen presenting cells (APC) to allow activation of "memory" T cells, as compared to "naive" T cells. At the end of the culture period, T cell activity is detected using assays for T cell activity including ^{51}Cr release involving peptide-sensitized targets, T cell proliferation, or lymphokine release.

The following describes the peptide epitopes and corresponding nucleic acids of the invention.

IV.C. Binding Affinity of Peptide Epitopes for HLA Molecules

The large degree of HLA polymorphism is an important factor to consider with the epitope-based approach to vaccine development. To address this factor, epitope selection including identification of peptides capable of binding at high or intermediate affinity to multiple HLA molecules is often utilized, most preferably these epitopes bind at high or intermediate affinity to two or more allele specific HLA molecules.

CTL-inducing peptides of interest for vaccine compositions preferably include those that have an IC_{50} or binding affinity value for class I HLA molecules of 500 nM or better (*i.e.*, the value is ≤ 500 nM). HTL-inducing peptides preferably include those that have an IC_{50} or binding affinity value for class II HLA molecules of 1000 nM or better, (*i.e.*, the value is $\leq 1,000$ nM). For example, peptide binding is assessed by testing the capacity of a candidate peptide to bind to a purified HLA molecule *in vitro*. Peptides exhibiting high or intermediate affinity are then considered for further analysis. Selected peptides are tested on other members of the supertype family. In preferred embodiments, peptides that exhibit cross-reactive binding are then used in vaccines or in cellular screening analyses.

Higher HLA binding affinity is typically correlated with greater immunogenicity. Greater immunogenicity can be manifested in several different ways. Immunogenicity corresponds to whether an immune response is elicited at all, and to the vigor of any particular response, as well as to the extent of a population in which a response is elicited. For example, a peptide might elicit an immune response in a diverse array of the population, yet in no instance produce a vigorous response. In accordance with these principles, close to 90% of high binding peptides have been found to be immunogenic, as contrasted with about 50% of the peptides which bind with intermediate affinity. Moreover, higher binding affinity peptides leads to more vigorous immunogenic responses. As a result, less peptide is required to elicit a similar biological effect if a high affinity binding peptide is used. Thus, in preferred embodiments of the invention, high affinity binding epitopes are particularly useful.

The relationship between binding affinity for HLA class I molecules and immunogenicity of discrete peptide epitopes on bound antigens has been determined for the first time in the art by the present inventors. The correlation between binding affinity and immunogenicity was analyzed in two different experimental approaches (*see, e.g.*, Sette, *et al.*, *J. Immunol.* 153:5586-5592, 1994). In the first approach, the

- immunogenicity of potential epitopes ranging in HLA binding affinity over a 10,000-fold range was analyzed in HLA-A*0201 transgenic mice. In the second approach, the antigenicity of approximately 100 different hepatitis B virus (HBV)-derived potential epitopes, all carrying A*0201 binding motifs, was assessed by using PBL from acute
- 5 hepatitis patients. Pursuant to these approaches, it was determined that an affinity threshold value of approximately 500 nM (preferably 50 nM or less) determines the capacity of a peptide epitope to elicit a CTL response. These data are true for class I binding affinity measurements for naturally processed peptides and for synthesized T cell epitopes. These data also indicate the important role of determinant selection in the
- 10 shaping of T cell responses (*see, e.g., Schaeffer et al. Proc. Natl. Acad. Sci. USA* 86:4649-4653, 1989).

- An affinity threshold associated with immunogenicity in the context of HLA class II DR molecules has also been delineated (*see, e.g., Southwood et al. J. Immunology* 160:3363-3373, 1998). In order to define a biologically significant threshold of DR
- 15 binding affinity, a database of the binding affinities of 32 DR-restricted epitopes for their restricting element (*i.e., the HLA molecule that binds the motif*) was compiled. In approximately half of the cases (15 of 32 epitopes), DR restriction was associated with high binding affinities, *i.e.* binding affinity values of 100 nM or less. In the other half of the cases (16 of 32), DR restriction was associated with intermediate affinity (binding
- 20 affinity values in the 100-1000 nM range). In only one of 32 cases was DR restriction associated with an IC_{50} of 1000 nM or greater. Thus, 1000 nM can be defined as an affinity threshold associated with immunogenicity in the context of DR molecules.

The binding affinity of peptides for HLA molecules can be determined as described in Example 1, below.

25

IV.D. Peptide Epitope Binding Motifs and Supermotifs

- In the past few years evidence has accumulated to demonstrate that a large fraction of HLA class I and class II molecules can be classified into a relatively few supertypes, each characterized by largely overlapping peptide binding repertoires, and
- 30 consensus structures of the main peptide binding pockets.

For HLA molecule pocket analyses, the residues comprising the B and F pockets of HLA class I molecules as described in crystallographic studies were analyzed (*see, e.g., Guo, H. C. et al., Nature* 360:364, 1992; Saper, M. A., Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991; Madden, D. R., Garboczi, D. N. and Wiley, D. C.,

Cell 75:693, 1993; Parham, P., Adams, E. J., and Arnett, K. L., *Immunol. Rev.* 143:141, 1995). In these analyses, residues 9, 45, 63, 66, 67, 70, and 99 were considered to make up the B pocket; and the B pocket was deemed to determine the specificity for the amino acid residue in the second position of peptide ligands. Similarly, residues 77, 80, 81, and 5 116 were considered to determine the specificity of the F pocket; the F pocket was deemed to determine the specificity for the C-terminal residue of a peptide ligand bound by the HLA class I molecule.

Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues 10 required for allele-specific binding to HLA molecules have been identified. The presence of these residues correlates with binding affinity for HLA molecules. The identification of motifs and/or supermotifs that correlate with high and intermediate affinity binding is an important issue with respect to the identification of immunogenic peptide epitopes for the inclusion in a vaccine. Kast *et al.* (*J. Immunol.* 152:3904-3912, 1994) have shown 15 that motif-bearing peptides account for 90% of the epitopes that bind to allele-specific HLA class I molecules. In this study all possible peptides of 9 amino acids in length and overlapping by eight amino acids (240 peptides), which cover the entire sequence of the E6 and E7 proteins of human papillomavirus type 16, were evaluated for binding to five allele-specific HLA molecules that are expressed at high frequency among different 20 ethnic groups. This unbiased set of peptides allowed an evaluation of the predictive value of HLA class I motifs. From the set of 240 peptides, 22 peptides were identified that bound to an allele-specific HLA molecule with high or intermediate affinity. Of these 22 peptides, 20 (*i.e.* 91%) were motif-bearing. Thus, this study demonstrates the value of motifs for the identification of peptide epitopes for inclusion in a vaccine: application of 25 motif-based identification techniques eliminates screening of 90% of the potential epitopes in a target antigen protein sequence.

Such peptide epitopes are identified in the Tables described below.

Peptides of the present invention may also comprise epitopes that bind to MHC class II DR molecules. A greater degree of heterogeneity in both size and binding frame 30 position of the motif, relative to the N and C termini of the peptide, exists for class II peptide ligands. This increased heterogeneity of HLA class II peptide ligands is due to the structure of the binding groove of the HLA class II molecule which, unlike its class I counterpart, is open at both ends. Crystallographic analysis of HLA class II DRB*0101-peptide complexes showed that the major energy of binding is contributed by peptide

residues complexed with complementary pockets on the DRB*0101 molecules. An important anchor residue engages the deepest hydrophobic pocket (*see, e.g.,* Madden, D.R. *Ann. Rev. Immunol.* 13:587, 1995) and is referred to as position 1 (P1). P1 may represent the N-terminal residue of a class II binding peptide epitope, but more typically is flanked towards the N-terminus by one or more residues. Other studies have also pointed to an important role for the peptide residue in the 6th position towards the C-terminus, relative to P1, for binding to various DR molecules.

Thus, peptides of the present invention are identified by any one of several HLA-specific amino acid motifs (*see, e.g.,* Tables I-III). If the presence of the motif corresponds to the ability to bind several allele-specific HLA antigens, it is referred to as a supermotif. The HLA molecules that bind to peptides that possess a particular amino acid supermotif are collectively referred to as an HLA "supertype."

The peptide motifs and supermotifs described below, and summarized in Tables I-III, provide guidance for the identification and use of peptide epitopes in accordance with the invention.

Examples of peptide epitopes bearing a respective supermotif or motif are included in Tables as designated in the description of each motif or supermotif below. The Tables include a binding affinity ratio listing for some of the peptide epitopes. The ratio may be converted to IC₅₀ by using the following formula: IC₅₀ of the standard peptide/ratio = IC₅₀ of the test peptide (*i.e.,* the peptide epitope). The IC₅₀ values of standard peptides used to determine binding affinities for Class I peptides are shown in Table IV. The IC₅₀ values of standard peptides used to determine binding affinities for Class II peptides are shown in Table V. The peptides used as standards for the binding assays described herein are examples of standards; alternative standard peptides can also be used when performing such an analysis.

To obtain the peptide epitope sequences listed in each Table, protein sequence data from fourteen HCV isolates were evaluated for the presence of the designated supermotif or motif. The fourteen strains include HPCCGAA, HPCPLYPRE, HCV-HCMR, HCV-J1, HPCGENANTI, HPCGENOM, HPCHUMR, HPCJCG, HPCJTA, HCV-J483, HCV-JK1, HCV-N, HPCPOLP, and HCV-J8. Peptide epitopes were additionally evaluated on the basis of their conservancy among these fourteen strains. A criterion for conservancy requires that the entire sequence of an HLA class I binding peptide be totally conserved in 79% of the sequences available for a specific protein. Similarly, a criterion for conservancy requires that the entire 9-mer core region of an HLA class II binding

peptide be totally conserved in 79% of the sequences available for a specific protein. The percent conservancy of the selected peptide epitopes is indicated on the Tables. The frequency, *i.e.* the number of strains of the fourteen strains in which the totally conserved peptide sequence was identified, is also shown. The "position" column in the Tables designates the amino acid position of the HCV polyprotein that corresponds to the first amino acid residue of the epitope. The "number of amino acids" indicates the number of residues in the epitope sequence.

HLA Class I Motifs Indicative of CTL Inducing Peptide Epitopes:

The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs delineated below are summarized in Table I. The HLA class I motifs set out in Table I(a) are those most particularly relevant to the invention claimed here. Primary and secondary anchor positions are summarized in Table II. Allele-specific HLA molecules that comprise HLA class I supertype families are listed in Table VI.

IV.D.1. HLA-A1 supermotif

The HLA-A1 supermotif is characterized by the presence in peptide ligands of a small (T or S) or hydrophobic (L, I, V, or M) primary anchor residue in position 2, and an aromatic (Y, F, or W) primary anchor residue at the C-terminal position of the epitope.

The corresponding family of HLA molecules that bind to the A1 supermotif (*i.e.*, the HLA-A1 supertype) includes at least A*0101, A*2601, A*2602, A*2501, and A*3201 (*see, e.g.*, DiBrino, M. *et al.*, *J. Immunol.* 151:5930, 1993; DiBrino, M. *et al.*, *J. Immunol.* 152:620, 1994; Kondo, A. *et al.*, *Immunogenetics* 45:249, 1997). Other allele-specific HLA molecules predicted to be members of the A1 superfamily are shown in Table VI. Peptides binding to each of the individual HLA proteins can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the A1 supermotif are set forth in Table VII.

IV.D.2. HLA-A2 supermotif

Primary anchor specificities for allele-specific HLA-A2.1 molecules (Falk *et al.*, *Nature* 351:290-296, 1991; Hunt *et al.*, *Science* 255:1261-1263, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992) and cross-reactive binding within the HLA A2 family (Fruci *et al.*, *Human Immunol.* 38:187-192, 1993; Tanigaki *et al.*, *Human Immunol.*

39:155-162, 1994) have been described. The present inventors have defined additional primary anchor residues that determine cross-reactive binding to multiple allele-specific HLA A2 molecules (Ruppert *et al.*, *Cell* 74:929-937, 1993; Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). The HLA-A2 supermotif comprises peptide ligands with L, I, V, M, A, T, or Q as a primary anchor residue at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope.

The corresponding family of HLA molecules (*i.e.*, the HLA-A2 supertype that binds these peptides) is comprised of at least: A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, and A*6901. Other allele-specific HLA molecules predicted to be members of the A2 superfamily are shown in Table VI. As explained in detail below, binding to each of the individual allele-specific HLA molecules can be modulated by substitutions at the primary anchor and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise an A2 supermotif are set forth in Table VIII. The motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

IV.D.3. HLA-A3 supermotif

The HLA-A3 supermotif is characterized by the presence in peptide ligands of A, L, I, V, M, S, or, T as a primary anchor at position 2, and a positively charged residue, R or K, at the C-terminal position of the epitope (*e.g.*, in position 9 of 9-mers). Exemplary members of the corresponding family of HLA molecules (the HLA-A3 supertype) that bind the A3 supermotif include at least A*0301, A*1101, A*3101, A*3301, and A*6801. Other allele-specific HLA molecules predicted to be members of the A3 superfamily are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions of amino acids at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the A3 supermotif are set forth in Table IX.

IV.D.4. HLA-A24 supermotif

The HLA-A24 supermotif is characterized by the presence in peptide ligands of an aromatic (F, W, or Y) or hydrophobic aliphatic (L, I, V, M, or T) residue as a primary anchor in position 2, and Y, F, W, L, I, or M as primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind to the A24 supermotif (*i.e.*, the A24 supertype) includes at least A*2402, A*3001, and A*2301. Other allele-specific HLA molecules predicted to be members of the A24 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the A24 supermotif are set forth in Table X.

IV.D.5. HLA-B7 supermotif

The HLA-B7 supermotif is characterized by peptides bearing proline in position 2 as a primary anchor, and a hydrophobic or aliphatic amino acid (L, I, V, M, A, F, W, or Y) as the primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind the B7 supermotif (*i.e.*, the HLA-B7 supertype) is comprised of at least twenty six HLA-B proteins including: B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, and B*7801 (*see, e.g.*, Sidney, *et al.*, *J. Immunol.* 154:247, 1995; Barber, *et al.*, *Curr. Biol.* 5:179, 1995; Hill, *et al.*, *Nature* 360:434, 1992; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995). Other allele-specific HLA molecules predicted to be members of the B7 superfamily are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B7 supermotif are set forth in Table XI.

IV.D.6. HLA-B27 supermotif

The HLA-B27 supermotif is characterized by the presence in peptide ligands of a positively charged (R, H, or K) residue as a primary anchor at position 2, and a hydrophobic (F, Y, L, W, M, I, A, or V) residue as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA

molecules that bind to the B27 supermotif (*i.e.*, the B27 supertype) include at least B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, and B*7301. Other allele-specific HLA molecules predicted to be members of the B27 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B27 supermotif are set forth in Table XII.

10 IV.D.7. HLA-B44 supermotif

The HLA-B44 supermotif is characterized by the presence in peptide ligands of negatively charged (D or E) residues as a primary anchor in position 2, and hydrophobic residues (F, W, Y, L, I, M, V, or A) as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B44 supermotif (*i.e.*, the B44 supertype) include at least: B*1801, B*1802, B*3701, B*4001, B*4002, B*4006, B*4402, B*4403, and B*4006. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the supermotif.

20

IV.D.8. HLA-B58 supermotif

The HLA-B58 supermotif is characterized by the presence in peptide ligands of a small aliphatic residue (A, S, or T) as a primary anchor residue at position 2, and an aromatic or hydrophobic residue (F, W, Y, L, I, V, M, or A) as a primary anchor residue at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B58 supermotif (*i.e.*, the B58 supertype) include at least: B*1516, B*1517, B*5701, B*5702, and B*5801. Other allele-specific HLA molecules predicted to be members of the B58 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B58 supermotif are set forth in Table XIII.

IV.D.9. HLA-B62 supermotif

The HLA-B62 supermotif is characterized by the presence in peptide ligands of the polar aliphatic residue Q or a hydrophobic aliphatic residue (L, V, M, I, or P) as a primary anchor in position 2, and a hydrophobic residue (F, W, Y, M, I, V, L, or A) as a primary anchor at the C-terminal position of the epitope. Exemplary members of the corresponding family of HLA molecules that bind to the B62 supermotif (*i.e.*, the B62 supertype) include at least: B*1501, B*1502, B*1513, and B5201. Other allele-specific HLA molecules predicted to be members of the B62 superfamily are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Peptide epitopes that comprise the B62 supermotif are set forth in Table XIV.

IV.D.10. HLA-A1 motif

The HLA-A1 motif is characterized by the presence in peptide ligands of T, S, or M as a primary anchor residue at position 2 and the presence of Y as a primary anchor residue at the C-terminal position of the epitope. An alternative allele-specific A1 motif is characterized by a primary anchor residue at position 3 rather than position 2. This motif is characterized by the presence of D, E, A, or S as a primary anchor residue in position 3, and a Y as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA A1 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Peptide epitopes that comprise either A1 motif are set forth in Table XV. The epitopes comprising T, S, or M at position 2 and Y at the C-terminal position are also included in the listing of HLA-A1 supermotif-bearing peptide epitopes listed in Table VII.

IV.D.11. HLA-A*0201 motif

An HLA-A2*0201 motif was first determined to be characterized by the presence in peptide ligands of L or M as a primary anchor residue in position 2, and L or V as a primary anchor residue at the C-terminal position of a 9-residue peptide (Falk *et al.*, *Nature* 351:290-296, 1991). The A*0201 motif was also determined to further comprise an I at position 2 and I or A at the C-terminal position of a nine amino acid peptide (Hunt

- et al.*, Science 255:1261-1263, March 6, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992). Subsequently, the A*0201 allele-specific motif has been defined by the present inventors to additionally comprise V, A, T, or Q as a primary anchor residue at position 2, and M as a primary anchor residue at the C-terminal position of the epitope.
- 5 Additionally, the A*0201 allele-specific motif has been found to comprise a T at the C-terminal position (Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). Thus, the HLA-A*0201 motif comprises peptide ligands with L, I, V, M, A, T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope. The preferred and tolerated residues that characterize the
- 10 primary anchor positions of the HLA-A*0201 motif are identical to the residues describing the A2 supermotif. (For reviews of relevant data, *see, e.g.*, Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Ruppert *et al.*, *Cell* 74:929-937, 1993; Sidney *et al.*, *Immunol. Today* 17:261-266, 1996; Sette and Sidney, *Curr. Opin. in Immunol.* 10:478-482, 1998). Secondary anchor residues that characterize the A*0201 motif have
- 15 additionally been defined as disclosed herein. These are disclosed in Table II. Peptide binding to HLA-A*0201 molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

- Peptide epitopes that comprise an A*0201 motif are set forth in Table VIII. The
- 20 A*0201 motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

IV.D.12. HLA-A3 motif

- 25 The HLA-A3 motif is characterized by the presence in peptide ligands of L, M, V, I, S, A, T, F, C, G, or D as a primary anchor residue at position 2, and the presence of K, Y, R, H, F, or A as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the
- 30 motif.

The A3 supermotif primary anchor residues comprise a subset of the A3- and A11-allele specific motif primary anchor residues. Peptide epitopes that comprise the A3 motif are set forth in Table XVI. Those peptide epitopes that also comprise the A3 supermotif are also listed in Table IX.

IV.D.13. HLA-A11 motif

The HLA-A11 motif is characterized by the presence in peptide ligands of V, T, M, L, I, S, A, G, N, C, D, or F as a primary anchor residue in position 2, and K, R, Y, or H as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A11 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Peptide epitopes that comprise the A11 motif are set forth in Table XVII; peptide epitopes comprising the A3 allele-specific motif are also present in this Table because of the overlap between the A3 and A11 motif primary anchor specificities. Further, those peptide epitopes that comprise the A3 supermotif are also listed in Table IX.

IV.D.14. HLA-A24 motif

The HLA-A24 motif is characterized by the presence in peptide ligands of Y, F, W, or M as a primary anchor residue in position 2, and F, L, I, or W as a primary anchor residue at the C-terminal position of the epitope. Peptide binding to HLA-A24 molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the motif.

Peptide epitopes that comprise the A24 motif are set forth in Table XVIII. These epitopes are also listed in Table X, which sets forth HLA-A24-supermotif-bearing peptide epitopes, as the primary anchor residues characterizing the A24 allele-specific motif comprise a subset of the A24 supermotif primary anchor residues.

HLA Class II Binding Motifs

The primary and secondary anchor residues of the HLA class II peptide epitope supermotifs and motifs delineated below are summarized in Table III.

IV.D.15. HLA DR-1-4-7 supermotif

Motifs have also been identified for peptides that bind to three common HLA class II allele-specific HLA molecules: HLA DRB1*0401, DRB1*0101, and DRB1*0701. Collectively, the common residues from these motifs delineate the HLA DR-1-4-7 supermotif. Peptides that bind to these DR molecules carry a supermotif characterized by a large aromatic or hydrophobic residue (Y, F, W, L, I, V, or M) as a primary anchor residue in position 1, and a small, non-charged residue (S, T, C, A, P, V,

I, L, or M) as a primary anchor residue in position 6 of a 9-mer core region. Allele-specific secondary effects and secondary anchors for each of these HLA types have also been identified. These are set forth in Table III. Peptide binding to HLA- DRB1*0401, DRB1*0101, and/or DRB1*0701 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Conserved peptide epitopes *i.e.*, conserved in $\geq 79\%$ ($\geq 11/14$) of the HCV strains used for the present analysis, may be described as corresponding to epitopes containing a nine residue core comprising the DR-1-4-7 supermotif, and in which the 9 residue core is conserved in $\geq 79\%$ (wherein position 1 of the motif is at position 1 of the nine residue core). Conserved 9-mer core regions are set forth in Table XIXa. Respective exemplary peptide epitopes of 15 amino acid residues in length, each of which comprise a conserved nine residue core, are also shown in section "a" of the Table. Cross-reactive binding data for exemplary 15-residue supermotif-bearing peptides are shown in Table XIXb.

IV.D.16. HLA DR3 motifs

Two alternative motifs (*i.e.*, submotifs) characterize peptide epitopes that bind to HLA-DR3 molecules. In the first motif (submotif DR3A) a large, hydrophobic residue (L, I, V, M, F, or Y) is present in anchor position 1 of a 9-mer core, and D is present as an anchor at position 4, towards the carboxyl terminus of the epitope. As in other class II motifs, core position 1 may or may not occupy the peptide N-terminal position.

The alternative DR3 submotif provides for lack of the large, hydrophobic residue at anchor position 1, and/or lack of the negatively charged or amide-like anchor residue at position 4, by the presence of a positive charge at position 6 towards the carboxyl terminus of the epitope. Thus, for the alternative allele-specific DR3 motif (submotif DR3B): L, I, V, M, F, Y, A, or Y is present at anchor position 1; D, N, Q, E, S, or T is present at anchor position 4; and K, R, or H is present at anchor position 6. Peptide binding to HLA-DR3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Conserved 9-mer core regions (*i.e.*, those sequences that are conserved in at least 79% of the 14 HCV strains used for the analysis) corresponding to a nine residue sequence comprising the DR3A submotif (wherein position 1 of the motif is at position 1 of the nine residue core) are set forth in Table XXa. Respective exemplary peptide

epitopes of 15 amino acid residues in length, each of which comprise a conserved nine residue core, are also shown in Table XXa. Table XXb shows binding data of exemplary DR3 submotif A-bearing peptides.

Conserved 9-mer core regions (*i.e.*, those that are at least 79% conserved in the 14 HCV strains used for the analysis) comprising the DR3B submotif and respective exemplary 15-mer peptides comprising the DR3 submotif-B epitope are set forth in Table XXc. Table XXd shows binding data of exemplary DR3 submotif B-bearing peptides.

Each of the HLA class I or class II peptide epitopes set out in the Tables herein are deemed singly to be an inventive aspect of this application. Further, it is also an inventive aspect of this application that each peptide epitope may be used in combination with any other peptide epitope.

IV.E. Enhancing Population Coverage of the Vaccine

Vaccines that have broad population coverage are preferred because they are more commercially viable and generally applicable to the most people. Broad population coverage can be obtained using the peptides of the invention (and nucleic acid compositions that encode such peptides) through selecting peptide epitopes that bind to HLA alleles which, when considered in total, are present in most of the population. Table XXI lists the overall frequencies of the HLA class I supertypes in various ethnicities (Table XXIa) and the combined population coverage achieved by the A2-, A3-, and B7-supertypes (Table XXIb). The A2-, A3-, and B7 supertypes are each present on the average of over 40% in each of these five major ethnic groups. Coverage in excess of 80% is achieved with a combination of these supermotifs. These results suggest that effective and non-ethnically biased population coverage is achieved upon use of a limited number of cross-reactive peptides. Although the population coverage reached with these three main peptide specificities is high, coverage can be expanded to reach 95% population coverage and above, and more easily achieve truly multispecific responses upon use of additional supermotif or allele-specific motif bearing peptides.

The B44-, A1-, and A24-supertypes are present, on average, in a range from 25% to 40% in these major ethnic populations (Table XXIa). While less prevalent overall, the B27-, B58-, and B62 supertypes are each present with a frequency >25% in at least one major ethnic group (Table XXIa). Table XXIb summarizes the estimated prevalence of combinations of HLA supertypes that have been identified in five major ethnic groups.

The incremental coverage obtained by the inclusion of A1,- A24-, and B44-supertypes to the A2, A3, and B7 coverage, or all of the supertypes described herein, is shown.

- 5 The data presented herein, together with the previous definition of the A2-, A3-, and B7-supertypes, indicates that all antigens, with the possible exception of A29, B8, and B46, can be classified into a total of nine HLA supertypes. By including epitopes from the six most frequent supertypes, an average population coverage of 99% is obtained for five major ethnic groups..

IV.F. Immune Response-Stimulating Peptide Analogs

- 10 In general, CTL and HTL responses are not directed against all possible epitopes. Rather, they are restricted to a few "immunodominant" determinants (Zinkernagel, *et al.*, *Adv. Immunol.* 27:5159, 1979; Bennis, *et al.*, *J. Exp. Med.* 168:19351939, 1988; Rawle, *et al.*, *J. Immunol.* 146:3977-3984, 1991). It has been recognized that immunodominance (Benacerraf, *et al.*, *Science* 175:273-279, 1972) could be explained by either the ability of
- 15 a given epitope to selectively bind a particular HLA protein (determinant selection theory) (Vitiello, *et al.*, *J. Immunol.* 131:1635, 1983); Rosenthal, *et al.*, *Nature* 267:156-158, 1977), or to be selectively recognized by the existing TCR (T cell receptor) specificities (repertoire theory) (Klein, J., *IMMUNOLOGY, THE SCIENCE OF SELF/NONSELF DISCRIMINATION*, John Wiley & Sons, New York, pp. 270-310, 1982). It has been
- 20 demonstrated that additional factors, mostly linked to processing events, can also play a key role in dictating, beyond strict immunogenicity, which of the many potential determinants will be presented as immunodominant (Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993).

- The concept of dominance and subdominance is relevant to immunotherapy of
- 25 both infectious diseases and cancer. For example, in the course of chronic viral disease, recruitment of subdominant epitopes can be important for successful clearance of the infection, especially if dominant CTL or HTL specificities have been inactivated by functional tolerance, suppression, mutation of viruses and other mechanisms (Franco, *et al.*, *Curr. Opin. Immunol.* 7:524-531, 1995). In the case of cancer and tumor antigens,
- 30 CTLs recognizing at least some of the highest binding affinity peptides might be functionally inactivated. Lower binding affinity peptides are preferentially recognized at these times, and may therefore be preferred in therapeutic or prophylactic anti-cancer vaccines.

In particular, it has been noted that a significant number of epitopes derived from known non-viral tumor associated antigens (TAA) bind HLA class I with intermediate affinity (IC_{50} in the 50-500 nM range). For example, it has been found that 8 of 15 known TAA peptides recognized by tumor infiltrating lymphocytes (TIL) or CTL bound
5 in the 50-500 nM range. (These data are in contrast with estimates that 90% of known viral antigens were bound by HLA class I molecules with IC_{50} of 50 nM or less, while only approximately 10% bound in the 50-500 nM range (Sette, *et al.*, *J. Immunol.*, 153:558-5592, 1994). In the cancer setting this phenomenon is probably due to elimination or functional inhibition of the CTL recognizing several of the highest binding
10 peptides, presumably because of T cell tolerization events.

Without intending to be bound by theory, it is believed that because T cells to dominant epitopes may have been clonally deleted, selecting subdominant epitopes may allow existing T cells to be recruited, which will then lead to a therapeutic or prophylactic response. However, the binding of HLA molecules to subdominant epitopes is often less
15 vigorous than to dominant ones. Accordingly, there is a need to be able to modulate the binding affinity of particular immunogenic epitopes for one or more HLA molecules, and thereby to modulate the immune response elicited by the peptide, for example to prepare analog peptides which elicit a more vigorous response. This ability would greatly enhance the usefulness of peptide-based vaccines and therapeutic agents.

Although peptides with suitable cross-reactivity among all alleles of a superfamily are identified by the screening procedures described above, cross-reactivity is not always as complete as possible, and in certain cases procedures to increase cross-reactivity of peptides can be useful; moreover, such procedures can also be used to modify other properties of the peptides such as binding affinity or peptide stability. Having established
25 the general rules that govern cross-reactivity of peptides for HLA alleles within a given motif or supermotif, modification (*i.e.*, analoging) of the structure of peptides of particular interest in order to achieve broader (or otherwise modified) HLA binding capacity can be performed. More specifically, peptides which exhibit the broadest cross-reactivity patterns, can be produced in accordance with the teachings herein. The present
30 concepts related to analog generation are set forth in greater detail in co-pending U.S.S.N. 09/226,775 filed 1/6/99.

In brief, the strategy employed utilizes the motifs or supermotifs which correlate with binding to certain HLA molecules. The motifs or supermotifs are defined by having primary anchors, and in many cases secondary anchors. Analog peptides can be created

by substituting amino acid residues at primary anchor, secondary anchor, or at primary and secondary anchor positions. Generally, analogs are made for peptides that already bear a motif or supermotif. Preferred secondary anchor residues of supermotifs and motifs that have been defined for HLA class I and class II binding peptides are shown in
5 Tables II and III, respectively.

For a number of the motifs or supermotifs in accordance with the invention, residues are defined which are deleterious to binding to allele-specific HLA molecules or members of HLA supertypes that bind the respective motif or supermotif (Tables II and III). Accordingly, removal of such residues that are detrimental to binding can be
10 performed in accordance with the present invention. For example, in the case of the A3 supertype, when all peptides that have such deleterious residues are removed from the population of analyzed peptides, the incidence of cross-reactivity increases from 22% to 37% (see, e.g., Sidney, J. *et al.*, *Hu. Immunol.* 45:79, 1996). Thus, one strategy to improve the cross-reactivity of peptides within a given supermotif is simply to delete one
15 or more of the deleterious residues present within a peptide and substitute a small "neutral" residue such as Ala (that may not influence T cell recognition of the peptide). An enhanced likelihood of cross-reactivity is expected if, together with elimination of detrimental residues within a peptide, "preferred" residues associated with high affinity binding to an allele-specific HLA molecule or to multiple HLA molecules within a
20 superfamily are inserted.

To ensure that an analog peptide, when used as a vaccine, actually elicits a CTL response to the native epitope *in vivo* (or, in the case of class II epitopes, elicits helper T cells that cross-react with the wild type peptides), the analog peptide may be used to immunize T cells *in vitro* from individuals of the appropriate HLA allele. Thereafter, the
25 immunized cells' capacity to induce lysis of wild type peptide sensitized target cells is evaluated. It will be desirable to use as antigen presenting cells, cells that have been either infected, or transfected with the appropriate genes, or, in the case of class II epitopes only, cells that have been pulsed with whole protein antigens, to establish whether endogenously produced antigen is also recognized by the relevant T cells.

Another embodiment of the invention is to create analogs of weak binding peptides, to thereby ensure adequate numbers of cross-reactive cellular binders. Class I binding peptides exhibiting binding affinities of 500-5000 nM, and carrying an acceptable but suboptimal primary anchor residue at one or both positions can be "fixed" by

substituting preferred anchor residues in accordance with the respective supertype. The analog peptides can then be tested for crossbinding activity.

Another embodiment for generating effective peptide analogs involves the substitution of residues that have an adverse impact on peptide stability or solubility in, e.g., a liquid environment. This substitution may occur at any position of the peptide epitope. For example, a cysteine (C) can be substituted out in favor of α -amino butyric acid. Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substituting α -amino butyric acid for C not only alleviates this problem, but actually improves binding and crossbinding capability in certain instances (*see, e.g.*, the review by Sette *et al.*, In: Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999). Substitution of cysteine with α -amino butyric acid may occur at any residue of a peptide epitope, *i.e.* at either anchor or non-anchor positions.

Representative analog peptides are set forth in Table XXII. The Table indicates the length and sequence of the analog peptide as well as the motif or supermotif, if appropriate. The information in the "Fixed Nomenclature" column indicates the residues substituted at the indicated position numbers for the respective analog.

IV.G. Computer Screening of Protein Sequences from Disease-Related Antigens for Supermotif- or Motif-Bearing Peptides

In order to identify supermotif- or motif-bearing epitopes in a target antigen, a native protein sequence, e.g., a tumor-associated antigen, or sequences from an infectious organism, or a donor tissue for transplantation, is screened using a means for computing, such as an intellectual calculation or a computer, to determine the presence of a supermotif or motif within the sequence. The information obtained from the analysis of native peptide can be used directly to evaluate the status of the native peptide or may be utilized subsequently to generate the peptide epitope.

Computer programs that allow the rapid screening of protein sequences for the occurrence of the subject supermotifs or motifs are encompassed by the present invention; as are programs that permit the generation of analog peptides. These programs are implemented to analyze any identified amino acid sequence or operate on an unknown sequence and simultaneously determine the sequence and identify motif-bearing epitopes thereof; analogs can be simultaneously determined as well. Generally, the identified sequences will be from a pathogenic organism or a tumor-associated peptide. For

example, the target molecules considered herein include, without limitation, the core, S, E1, NS1/E2, NS2, NS3, NS4, and NS5 regions of HCV.

In cases where the sequence of multiple variants of the same target protein are available, peptides may also be selected on the basis of their conservancy. A presently preferred criterion for conservancy defines that the entire sequence of an HLA class I binding peptide or the entire 9-mer core of a class II binding peptide, be totally (*i.e.*, 100%) conserved in at least 79% of the sequences evaluated for a specific protein. This definition of conservancy has been employed herein; although, as appreciated by those in the art, lower or higher degrees of conservancy can be employed as appropriate for a given antigenic target.

It is important that the selection criteria utilized for prediction of peptide binding are as accurate as possible, to correlate most efficiently with actual binding. Prediction of peptides that bind, for example, to HLA-A*0201, on the basis of the presence of the appropriate primary anchors, is positive at about a 30% rate (see, *e.g.*, Ruppert, J. *et al. Cell* 74:929, 1993). However, by extensively analyzing peptide-HLA binding data disclosed herein, data in related patent applications, and data in the art, the present inventors have developed a number of allele-specific polynomial algorithms that dramatically increase the predictive value over identification on the basis of the presence of primary anchor residues alone. These algorithms take into account not only the presence or absence of primary anchors, but also consider the positive or deleterious presence of secondary anchor residues (to account for the impact of different amino acids at different positions). The algorithms are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA interactions can be approximated as a linear polynomial function of the type:

$$\Delta G = a_{j1} \times a_{j2} \times a_{j3} \dots \times a_{jn}$$

where a_{ji} is a coefficient that represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. An important assumption of this method is that the effects at each position are essentially independent of each other. This assumption is justified by studies that demonstrated that peptides are bound to HLA molecules and recognized by T cells in essentially an extended conformation. Derivation of specific algorithm coefficients has been described, for example, in Gulukota, K. *et al.*, *J. Mol. Biol.* 267:1258, 1997.

Additional methods to identify preferred peptide sequences, which also make use of specific motifs, include the use of neural networks and molecular modeling programs (see, e.g., Milik *et al.*, *Nature Biotechnology* 16:753, 1998; Altuvia *et al.*, *Hum. Immunol.* 58:1, 1997; Altuvia *et al.*, *J. Mol. Biol.* 249:244, 1995; Buus, S. *Curr. Opin. Immunol.* 11:209-213, 1999; Brusic, V. *et al.*, *Bioinformatics* 14:121-130, 1998; Parker *et al.*, *J. Immunol.* 152:163, 1993; Meister *et al.*, *Vaccine* 13:581, 1995; Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994; Sturniolo *et al.*, *Nature Biotechnol.* 17:555 1999).

For example, it has been shown that in sets of A*0201 motif-bearing peptides containing at least one preferred secondary anchor residue while avoiding the presence of any deleterious secondary anchor residues, 69% of the peptides will bind A*0201 with an IC₅₀ less than 500 nM (Ruppert, J. *et al. Cell* 74:929, 1993). These algorithms are also flexible in that cut-off scores may be adjusted to select sets of peptides with greater or lower predicted binding properties, as desired.

In utilizing computer screening to identify peptide epitopes, a protein sequence or translated sequence may be analyzed using software developed to search for motifs, for example the "FINDPATTERNS" program (Devereux, *et al. Nucl. Acids Res.* 12:387-395, 1984) or MotifSearch 1.4 software program (D. Brown, San Diego, CA) to identify potential peptide sequences containing appropriate HLA binding motifs. The identified peptides can be scored using customized polynomial algorithms to predict their capacity to bind specific HLA class I or class II alleles. As appreciated by one of ordinary skill in the art, a large array of computer programming software and hardware options are available in the relevant art which can be employed to implement the motifs of the invention in order to evaluate (e.g., without limitation, to identify epitopes, identify epitope concentration per peptide length, or to generate analogs) known or unknown peptide sequences.

In accordance with the procedures described above, HCV peptide epitopes and analogs thereof that are able to bind HLA supertype groups or allele-specific HLA molecules have been identified (Tables VII-XX; Table XXII).

IV.H. Preparation of Peptide Epitopes

Peptides in accordance with the invention can be prepared synthetically, by recombinant DNA technology or chemical synthesis, or from natural sources such as native tumors or pathogenic organisms. Peptide epitopes may be synthesized individually or as polypeptidic peptides. Although the peptide will preferably be substantially free of

other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides may be synthetically conjugated to native fragments or particles.

The peptides in accordance with the invention can be a variety of lengths, and either in their neutral (uncharged) forms or in forms which are salts. The peptides in accordance with the invention are either free of modifications such as glycosylation, side chain oxidation, or phosphorylation; or they contain these modifications, subject to the condition that modifications do not destroy the biological activity of the peptides as described herein.

The peptides of the invention can be prepared in a wide variety of ways. For the preferred relatively short size, the peptides can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. (See, for example, Stewart & Young, SOLID PHASE PEPTIDE SYNTHESIS, 2D. ED., Pierce Chemical Co., 1984). Further, individual peptide epitopes can be joined using chemical ligation to produce larger peptides that are still within the bounds of the invention.

Alternatively, recombinant DNA technology can be employed wherein a nucleotide sequence which encodes an immunogenic peptide of interest is inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). Thus, recombinant polypeptides which comprise one or more peptide sequences of the invention can be used to present the appropriate T cell epitope.

The nucleotide coding sequence for peptide epitopes of the preferred lengths contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al.*, *J. Am. Chem. Soc.* 103:3185 (1981). Peptide analogs can be made simply by substituting the appropriate and desired nucleic acid base(s) for those that encode the native peptide sequence; exemplary nucleic acid substitutions are those that encode an amino acid defined by the motifs/super motifs herein. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly available in the art, and the vectors used to transform suitable hosts to produce the desired fusion protein. A number of such vectors and suitable host systems are now available. For expression of the fusion proteins, the coding sequence will be provided with operably linked start and stop codons, promoter and

terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts. Of course, yeast, insect or mammalian cell hosts may also be used, employing suitable vectors and control sequences.

It is often preferable that the peptide epitope be as small as possible while still maintaining substantially all of the immunologic activity of the native protein. When possible, it may be desirable to optimize HLA class I binding peptide epitopes of the invention to a length of about 8 to about 13 amino acid residues, preferably 9 to 10. HLA class II binding peptide epitopes may be optimized to a length of about 6 to about 30 amino acids in length, preferably to between about 13 and about 20 residues. Preferably, the peptide epitopes are commensurate in size with endogenously processed pathogen-derived peptides or tumor cell peptides that are bound to the relevant HLA molecules, however, the identification and preparation of peptides of other lengths can also be carried out using the techniques described herein.

In alternative embodiments, peptides of the invention can be linked as a polypeptidic peptide, or as a minigene that encodes a polypeptidic peptide.

In another embodiment, it is preferred to identify native peptide regions that contain a high concentration of class I and/or class II epitopes. Such a sequence is generally selected on the basis that it contains the greatest number of epitopes per amino acid length. It is to be appreciated that epitopes can be present in a frame-shifted manner, e.g. a 10 amino acid long peptide could contain two 9 amino acid long epitopes and one 10 amino acid long epitope; upon intracellular processing, each epitope can be exposed and bound by an HLA molecule upon administration of such a peptide. This larger, preferably multi-epitopic, peptide can be generated synthetically, recombinantly, or via cleavage from the native source.

IV.I. Assays to Detect T-Cell Responses

Once HLA binding peptides are identified, they can be tested for the ability to elicit a T-cell response. The preparation and evaluation of motif-bearing peptides are described in PCT publications WO 94/20127 and WO 94/03205. Briefly, peptides comprising epitopes from a particular antigen are synthesized and tested for their ability to bind to the appropriate HLA proteins. These assays may involve evaluating the

binding of a peptide of the invention to purified HLA class I molecules in relation to the binding of a radioiodinated reference peptide. Alternatively, cells expressing empty class I molecules (*i.e.* lacking peptide therein) may be evaluated for peptide binding by immunofluorescent staining and flow microfluorimetry. Other assays that may be used to evaluate peptide binding include peptide-dependent class I assembly assays and/or the inhibition of CTL recognition by peptide competition. Those peptides that bind to the class I molecule, typically with an affinity of 500 nM or less, are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in vitro* or *in vivo* CTL responses that can give rise to CTL populations capable of reacting with selected target cells associated with a disease. Corresponding assays are used for evaluation of HLA class II binding peptides. HLA class II motif-bearing peptides that are shown to bind, typically at an affinity of 1000 nM or less, are further evaluated for the ability to stimulate HTL responses.

Conventional assays utilized to detect T cell responses include proliferation assays, lymphokine secretion assays, direct cytotoxicity assays, and limiting dilution assays. For example, antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells. Alternatively, mutant non-human mammalian cell lines that are deficient in their ability to load class I molecules with internally processed peptides and that have been transfected with the appropriate human class I gene, may be used to test for the capacity of the peptide to induce *in vitro* primary CTL responses.

Peripheral blood mononuclear cells (PBMCs) may be used as the responder cell source of CTL precursors. The appropriate antigen-presenting cells are incubated with peptide, after which the peptide-loaded antigen-presenting cells are then incubated with the responder cell population under optimized culture conditions. Positive CTL activation can be determined by assaying the culture for the presence of CTLs that kill radio-labeled target cells, both specific peptide-pulsed targets as well as target cells expressing endogenously processed forms of the antigen from which the peptide sequence was derived.

More recently, a method has been devised which allows direct quantification of antigen-specific T cells by staining with Fluorescein-labelled HLA tetrameric complexes (Altman, J. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:10330, 1993; Altman, J. D. *et al.*, *Science* 274:94, 1996). Other relatively recent technical developments include staining

for intracellular lymphokines, and interferon release assays or ELISPOT assays. Tetramer staining, intracellular lymphokine staining and ELISPOT assays all appear to be at least 10-fold more sensitive than more conventional assays (Lalvani, A. *et al.*, *J. Exp. Med.* 186:859, 1997; Dunbar, P. R. *et al.*, *Curr. Biol.* 8:413, 1998; Murali-Krishna, K. *et al.*, *Immunity* 8:177, 1998).

HTL activation may also be assessed using such techniques known to those in the art such as T cell proliferation and secretion of lymphokines, *e.g.* IL-2 (*see, e.g.* Alexander *et al.*, *Immunity* 1:751-761, 1994).

Alternatively, immunization of HLA transgenic mice can be used to determine immunogenicity of peptide epitopes. Several transgenic mouse models including mice with human A2.1, A11 (which can additionally be used to analyze HLA-A3 epitopes), and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed. Additional transgenic mouse models with other HLA alleles may be generated as necessary. Mice may be immunized with peptides emulsified in Incomplete Freund's Adjuvant and the resulting T cells tested for their capacity to recognize peptide-pulsed target cells and target cells transfected with appropriate genes. CTL responses may be analyzed using cytotoxicity assays described above. Similarly, HTL responses may be analyzed using such assays as T cell proliferation or secretion of lymphokines.

Exemplary immunogenic peptide epitopes are set out in Table XXIII.

IV.J. Use of Peptide Epitopes as Diagnostic Agents and for Evaluating Immune Responses

In one embodiment of the invention, HLA class I and class II binding peptides as described herein can be used as reagents to evaluate an immune response. The immune response to be evaluated can be induced by using as an immunogen any agent that may result in the production of antigen-specific CTLs or HTLs that recognize and bind to the peptide epitope(s) to be employed as the reagent. The peptide reagent need not be used as the immunogen. Assay systems that can be used for such an analysis include relatively recent technical developments such as tetramers, staining for intracellular lymphokines and interferon release assays, or ELISPOT assays.

For example, a peptide of the invention may be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of antigen-specific CTLs following exposure to a tumor cell antigen or an immunogen. The HLA-tetrameric

complex is used to directly visualize antigen-specific CTLs (*see, e.g., Ogg et al., Science* 279:2103-2106, 1998; and Altman *et al., Science* 174:94-96, 1996) and determine the frequency of the antigen-specific CTL population in a sample of peripheral blood mononuclear cells. A tetramer reagent using a peptide of the invention may be generated as follows: A peptide that binds to an HLA molecule is refolded in the presence of the corresponding HLA heavy chain and β_2 -microglobulin to generate a trimolecular complex. The complex is biotinylated at the carboxyl terminal end of the heavy chain at a site that was previously engineered into the protein. Tetramer formation is then induced by the addition of streptavidin. By means of fluorescently labeled streptavidin, the tetramer can be used to stain antigen-specific cells. The cells may then be identified, for example, by flow cytometry. Such an analysis may be used for diagnostic or prognostic purposes. Cells identified by the procedure can also be used for therapeutic purposes.

Peptides of the invention may also be used as reagents to evaluate immune recall responses. (*see, e.g., Bertoni et al., J. Clin. Invest.* 100:503-513, 1997 and Penna *et al., J. Exp. Med.* 174:1565-1570, 1991.) For example, patient PBMC samples from individuals with HCV infection may be analyzed for the presence of antigen-specific CTLs or HTLs using specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBMCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed, for example, for cytotoxic activity (CTL) or for HTL activity.

The peptides may also be used as reagents to evaluate the efficacy of a vaccine. PBMCs obtained from a patient vaccinated with an immunogen may be analyzed using, for example, either of the methods described above. The patient is HLA typed, and peptide epitope reagents that recognize the allele-specific molecules present in that patient are selected for the analysis. The immunogenicity of the vaccine is indicated by the presence of epitope-specific CTLs and/or HTLs in the PBMC sample.

The peptides of the invention may also be used to make antibodies, using techniques well known in the art (*see, e.g. CURRENT PROTOCOLS IN IMMUNOLOGY*, Wiley/Greene, NY; and *Antibodies A Laboratory Manual*, Harlow and Lane, Cold Spring Harbor Laboratory Press, 1989), which may be useful as reagents to diagnose or monitor cancer. Such antibodies include those that recognize a peptide in the context of an HLA molecule, *i.e.*, antibodies that bind to a peptide-MHC complex.

IV.K. Vaccine Compositions

- Vaccines and methods of preparing vaccines that contain an immunogenically effective amount of one or more peptides as described herein are further embodiments of the invention. Once appropriately immunogenic epitopes have been defined, they can be sorted and delivered by various means, herein referred to as "vaccine" compositions.
- Such vaccine compositions can include, for example, lipopeptides (e.g., Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (see, e.g., Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991; Alonso *et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995), peptide compositions contained in immune stimulating complexes (ISCOMS) (see, e.g., Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998), multiple antigen peptide systems (MAPs) (see e.g., Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996), viral delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990), particles of viral or synthetic origin (e.g., Kofler, N. *et al.*, *J. Immunol. Methods* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Faló, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993), liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996), or, naked or particle absorbed cDNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.
- Vaccines of the invention include nucleic acid-mediated modalities. DNA or RNA encoding one or more of the peptides of the invention can also be administered to a patient. This approach is described, for instance, in Wolff *et al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based

delivery technologies include "naked DNA", facilitated (bupivacaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687).

- For therapeutic or prophylactic immunization purposes, the peptides of the
- 5 invention can also be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. As an example of this approach, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host bearing a tumor, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits a host
- 10 CTL and/or HTL response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guérin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus
- 15 vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein.

- Furthermore, vaccines in accordance with the invention encompass compositions of one or more of the claimed peptide(s). A peptide can be present in a vaccine individually. Alternatively, the peptide can exist as a homopolymer comprising
- 20 multiple copies of the same peptide, or as a heteropolymer of various peptides. Polymers have the advantage of increased immunological reaction and, where different peptide epitopes are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the pathogenic organism or tumor-related peptide targeted for an immune response. The composition can be a
- 25 naturally occurring region of an antigen or can be prepared, *e.g.*, recombinantly or by chemical synthesis.

- Carriers that can be used with vaccines of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza, hepatitis B virus
- 30 core protein, and the like. The vaccines can contain a physiologically tolerable (*i.e.*, acceptable) diluent such as water, or saline, preferably phosphate buffered saline. The vaccines also typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by

conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerylcysteinylserine (P₃CSS).

Upon immunization with a peptide composition in accordance with the invention, via injection, aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing large amounts of CTLs and/or HTLs specific for the desired antigen. Consequently, the host becomes at least partially immune to later infection, or at least partially resistant to developing an ongoing chronic infection, or derives at least some therapeutic benefit when the antigen was tumor-associated.

In some embodiments it may be desirable to combine the class I peptide components with components that induce or facilitate neutralizing antibody responses to the target antigen of interest, particularly to viral envelope antigens. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. An alternative embodiment of such a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE™ (Epimmune, San Diego, CA) molecule (described, for example, in U.S. Patent Number 5,736,142).

A vaccine of the invention can also include antigen-presenting cells, such as dendritic cells, as a vehicle to present peptides of the invention. Vaccine compositions can be created *in vitro*, following dendritic cell mobilization and harvesting, whereby loading of dendritic cells occurs *in vitro*. For example, dendritic cells are transfected, e.g., with a minigene in accordance with the invention. The dendritic cell can then be administered to a patient to elicit immune responses *in vivo*.

Antigenic peptides are used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTL or HTL cells, can be used to treat tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a therapeutic vaccine peptide or nucleic acid in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular tumor-associated antigen are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction

(HTL) of their specific target cell (an infected cell or a tumor cell). Transfected dendritic cells may also be used as antigen presenting cells.

The vaccine compositions of the invention can also be used in combination with antiviral drugs such as interferon- α , or other treatments for viral infection.

- 5 Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polypeptidic composition for use in a vaccine, or for selecting discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. It is preferred that each of the following principles are balanced in order to make the selection. The multiple epitopes to be incorporated in a given vaccine
10 composition may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived.

Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polypeptidic composition for use in a vaccine, or for selecting discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. Exemplary epitopes that may be utilized in a vaccine to treat or prevent HCV
15 infection are set out in Tables XXVI-XXIX, and Table XXXII. It is preferred that each of the following principles are balanced in order to make the selection.

- 1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with HCV clearance. For HLA Class I
20 this includes 3-4 epitopes that come from at least one antigen of HCV. For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one HCV antigen (*see e.g.*, Rosenberg *et al.*, *Science* 278:1447-1450).

- 2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC_{50} of 500 nM or less, or for
25 Class II an IC_{50} of 1000 nM or less.

- 3.) Sufficient supermotif bearing-peptides, or a sufficient array of allele-specific motif-bearing peptides, are selected to give broad population coverage. For example, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth,
30 or redundancy of, population coverage.

- 4.) When selecting epitopes from cancer-related antigens it is often preferred to select analogs because the patient may have developed tolerance to the native epitope.

When selecting epitopes for infectious disease-related antigens it is preferable to select either native or analoged epitopes.

5.) Of particular relevance are epitopes referred to as "nested epitopes."

Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A nested peptide sequence can comprise both HLA class I and HLA class II epitopes.

When providing nested epitopes, it is preferable to provide a sequence that has the greatest number of epitopes per provided sequence. Preferably, one avoids providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a longer peptide sequence, such as a sequence comprising nested epitopes, it is important to screen the sequence in order to insure that it does not have pathological or other deleterious biological properties.

6.) If a polyepitopic protein is created, or when creating a minigene, an objective is to generate the smallest peptide that encompasses the epitopes of interest.

This principle is similar, if not the same as that employed when selecting a peptide comprising nested epitopes. However, with an artificial polyepitopic peptide, the size minimization objective is balanced against the need to integrate any spacer sequences between epitopes in the polyepitopic protein. Spacer amino acid residues can be introduced to avoid junctional epitopes (an epitope recognized by the immune system, not present in the target antigen, and only created by the man-made juxtaposition of epitopes), or to facilitate cleavage between epitopes and thereby enhance epitope presentation. Junctional epitopes are generally to be avoided because the recipient may generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

Examples of polyepitopic vaccine compositions designed based on the above criteria can include epitopes from the core, S, E1, NS1/E2, NS2, NS3, NS4, and NS5 domains of the HCV polyprotein. These regions encompass the following amino acid sequences using numbering relative to the prototype HCV-1 strain (Genbank accession number M62321; *see, e.g.*, US Patent Nos. 5,683,864 and 5,670,153): C domain (amino acids 1-120); S (amino acids 120-400); NS3 (amino acids 1050-1640); NS4 (amino acids 1640-2000); NS5 (amino acids 2000-3011); and envelop proteins, E1 and E2/NS1, encompassing amino acids 192-750. Amino acids 750 to 1050 are designated as domain X as applied to the present invention. As appreciated by one of ordinary skill in the art,

the designation of the amino acid range for each domain may diverge to some extent from that of HCV-1 depending on the strain of HCV. One of ordinary skill in the art, when looking at an HCV polyprotein sequence, would readily be able to determine the domain boundaries.

- 5 Specific embodiments of the polypeptopic compositions of the present invention include a pharmaceutical composition comprising a pharmaceutically acceptable carrier and combination of motif-bearing peptides that are immunologically cross-reactive with peptides of HCV-1, wherein at least one of the peptides bears a motif of Table Ia, and further wherein the combination of motif-bearing peptides consists of: a) one or more
10 peptides comprising at least 8 amino acids from an HCV C domain; b) one or more peptides comprising at least 8 amino acids of a further domain selected from the group consisting of: an S domain, an NS3 domain, an NS4 domain, or an NS5 domain, and; c) optionally, one or more motif-bearing peptides from one or more additional HCV domains with a *proviso* that an additional domain is not a further domain listed in "b".
- 15 Preferably, such a pharmaceutical composition may additionally comprise one or more distinct HCV motif-bearing peptide(s) comprising at least 8 amino acids of an X domain or, alternatively, the composition may further comprise additional HCV motif-bearing peptide(s) that are from an envelope domain, the envelope domain peptide(s) consisting of one or more copies of a single HCV peptide comprising at least 8 amino acids of an
20 envelope domain.

- In another embodiment, the polypeptopic pharmaceutical composition may comprise a pharmaceutically acceptable carrier and combination of motif-bearing peptides that are immunologically cross-reactive with HCV-1 peptides, the peptides from multiple domains of HCV, wherein at least one of the peptides bears a motif of Table Ia,
25 and wherein the combination of motif-bearing peptides consists essentially of: a) one or more peptides comprising at least 8 amino acids from a C domain; and, b) one or more peptides comprising at least 8 amino acids from an S, NS3, NS4, or NS5 domain, and, one HCV peptide comprising at least 8 amino acids of an envelope domain. Such a composition may further comprise one or more HCV motif-bearing peptides comprising
30 at least 8 amino acids of an X domain.

 Alternatively, a pharmaceutical composition of the invention may comprise: a) a pharmaceutically acceptable carrier; and, b) a combination of one or more motif-bearing peptides of at least 8 amino acids derived from one or more hepatitis C virus (HCV) domains, wherein said peptides are cross-reactive with peptides of HCV-1, with a *proviso*

that the combination does not include a peptide of at least 8 amino acids from an HCV C domain, and wherein at least one of the peptides bears a motif of Table Ia, said domains selected from the group consisting of: an S domain; an NS3 domain; an NS4 domain; an NS5 domain; and, an X domain. Such a composition may additionally comprise motif-bearing HCV envelope peptide(s) consisting of one or more copies of a single HCV peptide comprising at least 8 amino acids of an envelope domain.

- Lastly, an embodiment of the invention may comprise a pharmaceutical composition comprising a pharmaceutically acceptable carrier and combination of two or more motif-bearing peptides from a single domain of an HCV-1 strain, said peptides immunologically cross-reactive with peptides of an HCV-1 antigen, wherein at least one of the peptides bears a motif of Table Ia, and the peptides are derived from HCV, and the HCV domain is selected from the group consisting of: a C domain; an S domain; an NS3 domain; an NS4 domain; an NS5 domain; an X domain; or, an envelope domain from a single HCV strain, with a *proviso* that the envelope domain is other than a variable envelope domain.

In the embodiments set forth, "peptides immunologically cross-reactive with HCV-1" refers to peptides that are bound by the same antibody; "derived from" refers to a fragment or subsequence and conservatively modified variants thereof.

IV.K.1. Minigene Vaccines

- A number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding the peptides of the invention are a particularly useful embodiment of the invention. Epitopes for inclusion in a minigene are preferably selected according to the guidelines set forth in the previous section. A preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding a peptide comprising one or multiple epitopes of the invention.

- The use of multi-epitope minigenes is described below and in, *e.g.*, co-pending application U.S.S.N. 09/311,784; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid encoding supermotif- and/or motif-bearing HCV epitopes derived from multiple regions of the HCV polyprotein sequence, the PADRE™ universal helper T cell epitope (or

multiple HTL epitopes from HCV), and an endoplasmic reticulum-translocating signal sequence can be engineered.

The immunogenicity of a multi-epitopic minigene can be tested in transgenic mice to evaluate the magnitude of CTL induction responses against the epitopes tested.

- 5 Further, the immunogenicity of DNA-encoded epitopes *in vivo* can be correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. Thus, these experiments can show that the minigene serves to both: 1.) generate a CTL response and 2.) that the induced CTLs recognized cells expressing the encoded epitopes.

- 10 For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous polypeptide sequence is created. To optimize expression
15 and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequences that can be reverse translated and included in the minigene sequence include: HLA class I epitopes, HLA class II epitopes, a ubiquitination signal sequence, and/or an endoplasmic reticulum targeting signal. In addition, HLA presentation of CTL and HTL epitopes may be improved by including
20 synthetic (e.g. poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention.

- The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides
25 (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope polypeptide, can then be cloned into a desired expression vector.

- Standard regulatory sequences well known to those of skill in the art are
30 preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (e.g. ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, e.g., the human cytomegalovirus

(hCMV) promoter. See, e.g., U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence, if desired to enhance immunogenicity.

In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (included to enhance or decrease immunogenicity) can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (e.g., IL-2, IL-12, GM-CSF), cytokine-inducing molecules (e.g., LeIF), costimulatory molecules, or for HTL responses, pan-DR binding proteins (PADRE™, Epimmune, San Diego, CA). Helper (HTL) epitopes can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes; this allows direction of the HTL epitopes to a cell compartment different than that of the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the HLA class II pathway, thereby improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (e.g. TGF- β) may be beneficial in certain diseases.

Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and grown to saturation in shaker flasks or a bioreactor

according to well known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

- 5 Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene DNA vaccines, an alternative method for
- 10 formulating purified plasmid DNA may be desirable. A variety of methods have been described, and new techniques may become available. Cationic lipids can also be used in the formulation (see, e.g., as described by WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Pat No. 5,279,833; WO 91/06309; and Felgner, *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, glycolipids, fusogenic
- 15 liposomes, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

- Target cell sensitization can be used as a functional assay for expression and HLA
- 20 class I presentation of minigene-encoded CTL epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be
- 25 co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 (^{51}Cr) labeled and used as target cells for epitope specific CTL lines; cytotoxicity, detected by ^{51}Cr release, indicates both production of, and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL
- 30 activity.

In vivo immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (e.g., IM for DNA in PBS, intraperitoneal (IP) for lipid-complexed DNA).

Twenty-one days after immunization, splenocytes are harvested and restimulated for 1 week in the presence of peptides encoding each epitope being tested. Thereafter, for CTL effector cells, assays are conducted for cytolysis of peptide-loaded, ⁵¹Cr-labeled target cells using standard techniques. Lysis of target cells that were sensitized by HLA loaded with peptide epitopes, corresponding to minigene-encoded epitopes, demonstrates DNA vaccine function for *in vivo* induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment, DNA can be adhered to particles, such as gold particles.

IV.K.2. Combinations of CTL Peptides with Helper Peptides

Vaccine compositions comprising the peptides of the present invention, or analogs thereof, which have immunostimulatory activity may be modified to provide desired attributes, such as improved serum half life, or to enhance immunogenicity.

For instance, the ability of the peptides to induce CTL activity can be enhanced by linking the peptide to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. The use of T helper epitopes in conjunction with CTL epitopes to enhance immunogenicity is illustrated, for example, in co-pending U.S.S.N. 08/820360, U.S.S.N. 08/197,484, and U.S.S.N. 08/464,234.

Particularly preferred CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues. Alternatively, the CTL peptide may be linked to the T helper peptide without a spacer.

Although the CTL peptide epitope can be linked directly to the T helper peptide epitope, often CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological

conditions. The spacers are typically selected from, e.g., Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues. The CTL peptide epitope can be linked to the T helper peptide epitope either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the immunogenic peptide or the T helper peptide may be acylated.

HTL peptide epitopes can also be modified to alter their biological properties. For example, peptides comprising HTL epitopes can contain D-amino acids to increase their resistance to proteases and thus extend their serum half-life. Also, the epitope peptides of the invention can be conjugated to other molecules such as lipids, proteins or sugars, or any other synthetic compounds, to increase their biological activity. Specifically, the T helper peptide can be conjugated to one or more palmitic acid chains at either the amino or carboxyl termini.

In certain embodiments, the T helper peptide is one that is recognized by T helper cells present in the majority of the population. This can be accomplished by selecting amino acid sequences that bind to many, most, or all of the HLA class II molecules. These are known as "loosely HLA-restricted" or "promiscuous" T helper sequences. Examples of amino acid sequences that are promiscuous include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE), *Plasmodium falciparum* CS protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS), and Streptococcus 18kD protein at positions 116 (GAVDSILGGVATYGAA). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

Alternatively, it is possible to prepare synthetic peptides capable of stimulating T helper lymphocytes, in a loosely HLA-restricted fashion, using amino acid sequences not found in nature (see, e.g., PCT publication WO 95/07707). These synthetic compounds called Pan-DR-binding epitopes (e.g., PADRE™, Epimmune, Inc., San Diego, CA) are designed to most preferably bind most HLA-DR (human HLA class II) molecules. For instance, a pan-DR-binding epitope peptide having the formula: aKXVWANTLKAAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine, and a is either D-alanine or L-alanine, has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type.

An alternative of a pan-DR binding epitope comprises all "L" natural amino acids and can be provided in the form of nucleic acids that encode the epitope.

In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes cytotoxic T

- 5 lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo* against viral antigens. For example, palmitic acid residues can be attached to the ϵ - and α -amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser-, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle,
- 10 incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. In a preferred embodiment, a particularly effective immunogenic comprises palmitic acid attached to ϵ - and α - amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the immunogenic peptide.

- As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such
- 15 as tripalmitoyl-S-glycerylcysteinylserine (P₃CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide. (*See, e.g., Deres, et al., Nature* 342:561, 1989). Peptides of the invention can be coupled to P₃CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be
- 20 primed with P₃CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses to infection.

- As noted herein, additional amino acids can be added to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or
- 25 oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or the like, can be introduced at the C- or N-terminus of the peptide or oligopeptide, particularly class I peptides. However, it is to be noted that modification at the carboxyl terminus of a CTL epitope may, in some cases, alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the
- 30 natural sequence by being modified by terminal-NH₂ acylation, *e.g.*, by alkanoyl (C₁-C₂₀) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule.

Vaccine Compositions Comprising Dendritic Cells Pulsed with CTL and/or HTL Peptides

An embodiment of a vaccine composition in accordance with the invention comprises *ex vivo* administration of a cocktail of epitope-bearing peptides to PBMC, or isolated DC therefrom, from the patient's blood. A pharmaceutical to facilitate harvesting of DC can be used, such as GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides. In this embodiment, a vaccine comprises peptide-pulsed DCs which present the pulsed peptide epitopes complexed with HLA molecules on their surfaces. The vaccine is then administered to the patient.

IV.L. Administration of Vaccines for Therapeutic or Prophylactic Purposes

The peptides of the present invention and pharmaceutical and vaccine compositions of the invention are useful for administration to mammals, particularly humans, to treat and/or prevent HCV infection. Vaccine compositions containing the peptides of the invention are administered to a patient infected with HCV or to an individual susceptible to, or otherwise at risk for, HCV infection to elicit an immune response against HCV antigens and thus enhance the patient's own immune response capabilities. In therapeutic applications, peptide and/or nucleic acid compositions are administered to a patient in an amount sufficient to elicit an effective CTL and/or HTL response to the virus antigen and to cure or at least partially arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

The vaccine compositions of the invention may also be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 µg and the higher value is about 10,000; 20,000; 30,000; or 50,000 µg. Dosage values for a human typically range from about 500 µg to about 50,000 µg per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 µg to about 50,000 µg of peptide administered at defined intervals from about four weeks to six months after the

initial administration of vaccine. The immunogenicity of the vaccine may be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

- As noted above, peptides comprising CTL and/or HTL epitopes of the invention induce immune responses when presented by HLA molecules and contacted with a CTL or HTL specific for an epitope comprised by the peptide. The manner in which the peptide is contacted with the CTL or HTL is not critical to the invention. For instance, the peptide can be contacted with the CTL or HTL either *in vivo* or *in vitro*. If the contacting occurs *in vivo*, the peptide itself can be administered to the patient, or other vehicles, *e.g.*, DNA vectors encoding one or more peptides, viral vectors encoding the peptide(s), liposomes and the like, can be used, as described herein. When the peptide is contacted *in vitro*, the vaccinating agent can comprise a population of cells, *e.g.*, peptide-pulsed dendritic cells, or TAA-specific CTLs, which have been induced by pulsing antigen-presenting cells *in vitro* with the peptide. Such a cell population is subsequently administered to a patient in a therapeutically effective dose.

The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences.

- For pharmaceutical compositions, the immunogenic peptides of the invention, or DNA encoding them, are generally administered to an individual already infected with HCV. The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences. Those in the incubation phase or the acute phase of infection can be treated with the immunogenic peptides separately or in conjunction with other treatments, as appropriate.

- For therapeutic use, administration should generally begin at the first diagnosis of HCV infection. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. In chronic infection, loading doses followed by boosting doses may be required.

- Treatment of an infected individual with the compositions of the invention may hasten resolution of the infection in acutely infected individuals. For those individuals susceptible (or predisposed) to developing chronic infection, the compositions are particularly useful in methods for preventing the evolution from acute to chronic infection. Where susceptible individuals are identified prior to or during infection, the composition can be targeted to them, thus minimizing the need for administration to a larger population.

The peptide or other compositions used for the treatment or prophylaxis of HCV infection can be used, *e.g.*, in persons who have not manifested symptoms of disease but who act as a disease vector. In this context, it is generally important to provide an amount of the peptide epitope delivered by a mode of administration sufficient to effectively stimulate a cytotoxic T cell response; compositions which stimulate helper T cell responses can also be given in accordance with this embodiment of the invention.

The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 μg and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg . Dosage values for a human typically range from about 500 μg to about 50,000 μg per 70 kilogram patient. Boosting dosages of between about 1.0 μg to about 50000 μg of peptide pursuant to a boosting regimen over weeks to months may be administered depending upon the patient's response and condition as determined by measuring the specific activity of CTL and HTL obtained from the patient's blood. The peptides and compositions of the present invention may be employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides in preferred compositions of the invention, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

Thus, for treatment of chronic infection, a representative dose is in the range disclosed above, namely where the lower value is about 1, 5, 50, 500, or 1000 μg and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg , preferably from about 500 μg to about 50,000 μg per 70 kilogram patient. Initial doses followed by boosting doses at established intervals, *e.g.*, from four weeks to six months, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic infection, administration should continue until at least clinical symptoms or laboratory tests indicate that the viral infection has been eliminated or substantially abated and for a period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, intrathecal, or local administration. Preferably, the pharmaceutical compositions are administered parentally, *e.g.*, intravenously,

subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see, e.g., Remington's Pharmaceutical Sciences*, 17th Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985).

The peptides of the invention may also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or to target selectively to infected cells, as well as to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed

- from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing
- 5 liposomes, as described in, *e.g.*, Szoka, *et al.*, *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

- For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a
- 10 peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

- For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium
- 15 stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

- For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as
- 25 caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, *e.g.*, lecithin for intranasal
- 30 delivery.

IV.M. Kits

The peptide and nucleic acid compositions of this invention can be provided in kit form together with instructions for vaccine administration. Typically the kit would

- include desired peptide compositions in a container, preferably in unit dosage form and instructions for administration. An alternative kit would include a minigene construct with desired nucleic acids of the invention in a container, preferably in unit dosage form together with instructions for administration. Lymphokines such as IL-2 or IL-12 may also be included in the kit. Other kit components that may also be desirable include, for example, a sterile syringe, booster dosages, and other desired excipients.

- The invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of non-critical parameters that can be changed or modified to yield alternative embodiments in accordance with the invention.

V. EXAMPLES

- As in many viral diseases, there is evidence that clearance of HCV is mediated by CTL. In a study of primary HCV infection in six chimpanzees, four progressed to chronic infection (Cooper *et al.*, abstract, 19th US-Japan Hepatitis Joint Panel Meeting, January 27-29, 1998). It was found that these four animals showed either no CTL response or a very narrowly focused response during early infection. In contrast, in the remaining two animals that resolved the infection, a broad CTL response was observed against multiple HCV proteins, some of which were conserved. Weiner *et al.* (*Proc. Natl. Acad. Sci. USA* 92:2755-2759, 1995) demonstrated that viral escape, in which the epitopes presented to PATR class I molecules mutated, was linked with a progression toward chronic infection. These data show a role for the CTL in directing the course of HCV disease, and in shaping the genetic composition of HCV species in the persistently infected host.

- In work in humans, Koziel and co-workers have established the presence of HCV-specific CTL in liver infiltrates from patients with chronic HCV infection (Koziel *et al.*, *J. Immunol.* 149:3339, 1992; and Koziel *et al.*, *J. Virol.* 67:7522, 1993), and have also identified a number of CTL epitopes recognized in the context of several different HLA class I molecules. Other investigators have shown that HCV-specific CTL can be detected in the peripheral blood of patients with chronic hepatitis C (Cerny *et al.*, *J. Clin. Invest.* 95:521, 1995; Cerny *et al.*, *Curr. Topics in Micro. and Immunol* 189:169, 1994; Cerny *et al.*, Abst. 2nd International Meeting on Hepatitis C and Related Viruses; La Jolla, CA, 1994; Battegay *et al.*, Abst. 2nd International Meeting on Hepatitis C and Related

Viruses; La Jolla, CA, 1994; Shirai *et al.*, *J. Virol.* 68:3334, 1994; Shirai *et al.*, *J. Immunol.* 154:2733, 1995; Battegay *et al.*, *J. Virol.* 69:2462, 1995). In addition, escape variants have been demonstrated in patients chronically infected with HCV (Chang *et al.*, *J. Clin. Invest.* 100:2376-2385, 1997; Tsai *et al.*, *Gastroenterology* 115:954-966, 1998).

- 5 The magnitude of the CTL responses observed in HCV-infected patients is, in general, higher than those observed in the case of chronic hepatitis B infection, suggesting that there is less impairment of specific T cell immunity than with HBV infection. The magnitude of CTL responses in HCV patients is, however, lower than those observed in HBV infected individuals who successfully cleared HBV infection.
- 10 These results support the understanding that HCV infected patients are capable of responding to active immunotherapy, and suggest that potentiation and increasing of T cell responses to HCV may be of use in therapy and prevention of chronic HCV infection (Prince, A. M. *FEMS Micro. Rev.* 14:273, 1994).

- Several groups have analyzed the potential role of HCV-specific CTL responses
- 15 in disease resistance and pathogenesis. In some studies no correlation was found between CTL viremia and CTL precursor frequency for individual HCV epitopes (Rehermann *et al.*, *J. Clin. Invest.* 98:1432-1440, 1996; Wong *et al.*, *J. Immunol.* 160:1479-1488, 1998). In other studies, however, it was shown that a clear correlation existed between levels of HCV infection and CTL responses, provided that the global response against multiple
- 20 CTL epitopes was considered (Rehermann *et al.*, *J. Virol.* 70:7092-7102, 1996). These data represent a strong rationale for development of vaccine constructs capable of inducing vigorous CTL responses directed against a multiplicity of conserved HCV-derived epitopes.

- Koziel and colleagues have demonstrated the presence of HCV-specific CTLs, as
- 25 well as T helper cell responses, in exposed but seronegative individuals (Koziel *et al.*, *J. Infect. Diseases* 176:859-866, 1997). In addition, HCV-specific CTLs have been detected in healthy, seronegative family members of chronically HCV-infected patients, indicating that a protective immunity is established in absence of a detectable infection (Bronowicki *et al.*, *J. Infect. Dis.* 176:518-522, 1997; Scognamiglio *et al.*, in preparation).

- 30 Experimental evidence also indicates that HTL epitopes play an important role in immune reactivity and defenses against HCV infection (Missale *et al.*, *J. Clin. Invest.* 98:706-714, 1996). Diepolder *et al.* (in *Lancet* 346:1006, 1995) have shown that a region of the NS3 gene (NS3 1007-1534) is recognized by patients who clear acute HCV infection, but is not seen by patients who develop chronic infection. Subsequent studies

showed that this particular region contain a highly cross-reactive HTL epitope (NS3 1248-1261), which binds with good affinity to 10 of 13 DR molecules tested, and is highly conserved in 30/33 different HCV isolates considered (Diepolder *et al.*, *J. Virol.* 71:6011-6019, 1997). These data suggested that directing HTL responses to this type of epitope (rather than to less cross-reactive and/or highly variable ones) will be of therapeutic and prophylactic benefit and strongly argue for inclusion of this and other epitopes with similar characteristics in HCV vaccine constructs.

The following examples illustrate identification, selection, and use of immunogenic Class I and Class II peptide epitopes for inclusion in vaccine compositions.

Example 1: HLA Class I and Class II Binding Assays

The following example of peptide binding to HLA molecules demonstrates quantification of binding affinities of HLA class I and class II peptides. Binding assays can be performed with peptides that are either motif-bearing or not motif-bearing.

Epstein-Barr virus (EBV)-transformed homozygous cell lines, fibroblasts, CIR, or 721.22 transfectants were used as sources of HLA class I molecules. The specific cell lines routinely used for purification of MHC class I and class II molecules are listed in Table XXIV. Cell lysates were prepared and HLA molecules purified in accordance with disclosed protocols (Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). HLA molecules were purified from lysates by affinity chromatography. The lysate was passed over a column of Sepharose CL-4B beads coupled to an appropriate antibody. The antibodies used for the extraction of HLA from cell lysates are listed in Table XXV. The anti-HLA column was then washed with 10mM Tris-HCL, pH 8.0, in 1% NP-40, PBS, and PBS containing 0.4% n-octylglucoside and HLA molecules were eluted with 50mM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0. Eluates were then be concentrated by centrifugation in Centrprep 30 concentrators (Amicon, Beverly, MA). Protein content was evaluated by a BCA protein assay (Pierce Chemical Co., Rockford, IL) and confirmed by SDS-PAGE.

A detailed description of the protocol utilized to measure the binding of peptides to Class I and Class II MHC has been published (Sette *et al.*, *Mol. Immunol.* 31:813, 1994; Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998). Briefly, purified MHC molecules (5 to 500nM)

- were incubated with various unlabeled peptide inhibitors and 1-10nM ^{125}I -radiolabeled probe peptides for 48h in PBS containing 0.05% Nonidet P-40 (NP40) (or 20% w/v digitonin for H-2 IA assays) in the presence of a protease inhibitor cocktail. All assays were at pH 7.0 with the exception of DRB1*0301, which was performed at pH 4.5, and
- 5 DRB1*1601 (DR2w21 β_1) and DRB4*0101 (DRw53), which were performed at pH 5.0.

- Following incubation, MHC-peptide complexes were separated from free peptide by gel filtration on 7.8 mm x 15 cm TSK200 columns (TosoHaas 16215, Montgomeryville, PA). Because the large size of the radiolabeled peptide used for the DRB1*1501 (DR2w2 β_1) assay makes separation of bound from unbound peaks more
- 10 difficult under these conditions, all DRB1*1501 (DR2w2 β_1) assays were performed using a 7.8mm x 30cm TSK2000 column eluted at 0.6 mls/min. The eluate from the TSK columns was passed through a Beckman 170 radioisotope detector, and radioactivity was plotted and integrated using a Hewlett-Packard 3396A integrator, and the fraction of peptide bound was determined.

- 15 Radiolabeled peptides were iodinated using the chloramine-T method. Representative radiolabeled probe peptides utilized in each assay, and its assay specific IC_{50} nM, are summarized in Tables IV and V. Typically, in preliminary experiments, each MHC preparation was titrated in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of
- 20 the total radioactivity. All subsequent inhibition and direct binding assays were performed using these HLA concentrations.

- Since under these conditions $[\text{label}] < [\text{HLA}]$ and $\text{IC}_{50} \geq [\text{HLA}]$, the measured IC_{50} values are reasonable approximations of the true K_D values. Peptide inhibitors are typically tested at concentrations ranging from 120 $\mu\text{g/ml}$ to 1.2 ng/ml, and are tested in
- 25 two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the IC_{50} of a positive control for inhibition by the IC_{50} for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For database purposes, and inter-experiment comparisons, relative binding values are compiled. These values
- 30 can subsequently be converted back into IC_{50} nM values by dividing the IC_{50} nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation has proven to be the most accurate and consistent for

comparing peptides that have been tested on different days, or with different lots of purified MHC.

- Because the antibody used for HLA-DR purification (LB3.1) is α -chain specific, β_1 molecules are not separated from β_3 (and/or β_4 and β_5) molecules. The β_1 specificity of the binding assay is obvious in the cases of DRB1*0101 (DR1), DRB1*0802 (DR8w2), and DRB1*0803 (DR8w3), where no β_3 is expressed. It has also been demonstrated for DRB1*0301 (DR3) and DRB3*0101 (DR52a), DRB1*0401 (DR4w4), DRB1*0404 (DR4w14), DRB1*0405 (DR4w15), DRB1*1101 (DR5), DRB1*1201 (DR5w12), DRB1*1302 (DR6w19) and DRB1*0701 (DR7). The problem of β chain specificity for DRB1*1501 (DR2w2 β_1), DRB5*0101 (DR2w2 β_2), DRB1*1601 (DR2w21 β_1), DRB5*0201 (DR51Dw21), and DRB4*0101 (DRw53) assays is circumvented by the use of fibroblasts. Development and validation of assays with regard to DR β molecule specificity have been described previously (*see, e.g., Southwood et al., J. Immunol.* 160:3363-3373, 1998).
- Binding assays as outlined above may be used to analyze supermotif and/or motif-bearing epitopes as, for example, described in Example 2.

Example 2. Identification of Conserved HLA Supermotif- and Motif-Bearing CTL Candidate Epitopes

- Vaccine compositions of the invention may include multiple epitopes that comprise multiple HLA supermotifs or motifs to achieve broad population coverage. This example illustrates the identification of supermotif- and motif-bearing epitopes for the inclusion in such a vaccine composition. Calculation of population coverage was performed using the strategy described below.

Computer searches and algorithms for identification of supermotif and/or motif-bearing epitopes

- Computer searches for epitopes bearing HLA Class I or Class II supermotifs or motifs were performed as follows. All translated HCV isolate sequences were analyzed using a text string search software program, *e.g.,* MotifSearch 1.4 (D. Brown, San Diego) to identify potential peptide sequences containing appropriate HLA binding motifs; alternative programs are readily produced in accordance with information in the art in view of the motif/supermotif disclosure herein. Furthermore, such calculations can be

made mentally. Identified A2-, A3-, and DR-supermotif sequences were scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms take into account both extended and refined motifs (that is, to account for the impact of different amino acids at different positions), and are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$-\Delta G = a_{1i} \times a_{2i} \times a_{3i} \dots \times a_{ni}$$

where a_{ji} is a coefficient which represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue j occurs at position i in the peptide, it is assumed to contribute a constant amount j_i to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. This assumption is justified by studies from our laboratories that demonstrated that peptides are bound to MHC and recognized by T cells in essentially an extended conformation (data omitted herein).

The method of derivation of specific algorithm coefficients has been described in Gulukota *et al.*, *J. Mol. Biol.* 267:1258-126, 1997; (see also Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; and Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998).

Briefly, for all i positions, anchor and non-anchor alike, the geometric mean of the average relative binding (ARB) of all peptides carrying j is calculated relative to the remainder of the group, and used as the estimate of j_i . For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

Selection of HLA-A2 supertype cross-reactive peptides

Complete polyprotein sequences from fourteen HCV isolates were aligned, then scanned, utilizing motif identification software, to identify conserved 9- and 10-mer sequences containing the HLA-A2-supermotif main anchor specificity.

A total of 231 conserved, HLA-A2 supermotif-positive sequences were identified. These peptides were then evaluated for the presence of A*0201 preferred secondary anchor residues using A*0201-specific polynomial algorithms. A total of 67 conserved, motif-bearing and algorithm-positive sequences were identified.

5 Fifty of these conserved, motif-containing 9- and 10-mer peptides were tested for their capacity to bind to purified HLA-A*0201 molecules *in vitro* (HLA-A*0201 is considered a prototype A2 supertype molecule). Sixteen peptides bound A*0201 with IC₅₀ values ≤500 nM; 4 with high binding affinities (IC₅₀ values ≤50 nM) and 12 with intermediate binding affinities, in the 50-500 nM range (Table XXVI).

10 These 16 peptides were then tested for binding to additional A2-supertype molecules (A*0202, A*0203, A*0206, and A*6802). As shown in Table XXVI, most of these peptides were found to be A2-supertype cross-reactive binders. More specifically, 12/16 (75%) peptides bound at least three of the five A2-supertype molecules tested.

15 *Selection of HLA-A3 supermotif-bearing epitopes*

 The sequences from the same fourteen known HCV isolates scanned above were also examined for the presence of conserved peptides with the HLA-A3-supermotif primary anchors. A total of 71 conserved 9- or 10-mer motif containing sequences were identified. Further analysis using the A03 and A11 algorithms (see, e.g., Gulukota et al,
20 *J. Mol. Biol.* 267:1258-1267, 1997 and Sidney et al, *Human Immunol.* 45:79-93, 1996) identified 39 sequences that scored high in either or both algorithms. Twenty seven of the 39 peptides were synthesized and tested for binding to HLA-A*03 and HLA-A*11, the two most prevalent A3-supertype molecules. Fifteen peptides were identified which bound A3 and/or A11 with binding affinities of ≤500 nM (Table XXVII). These peptides
25 were then tested for binding cross-reactivity to the other common A3-supertype alleles (A*3101, A*3301, and A*6801). Seven of the 15 peptides bound at least three of the five HLA-A3-supertype molecules tested.

 In the course of an independent series of experiments (Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994), one peptide, HCV NS3 1262, not identified by the selection
30 criteria utilized above because it does not have the A3-supermotif main anchor specificity, was determined to be cross-reactive in the A3-supertype, binding A*03, A*11, and A*6801. It is also shown in Table XXVII. Interestingly, this peptide

represents a single residue N-terminal truncation of peptide 1073.14, which is also shown in Table XXVII.

In summary, 8 peptides that bind 3 or more A3-supertype molecules derived from conserved regions of the HCV genome were identified.

Selection of HLA-B7 supermotif bearing epitopes

When the same fourteen HCV isolates were also analyzed for the presence of conserved 9- or 10-mer peptides with the HLA-B7-supermotif, 35 sequences were identified. The corresponding peptides were synthesized and tested for binding to HLA-

B*0702, the most common B7-supertype allele (*i.e.*, the prototype B7 supertype allele). Thirteen peptides bound B*0702 with IC_{50} of ≤ 500 nM (Table XXVIIIa). These 13 peptides were then tested for binding to other common B7-supertype molecules (B*3501, B*51, B*5301, and B*5401). As shown in Table XXVIIIa, only 1 peptide (Core 169) was capable of binding to three or more of the five B7-supertype alleles tested.

To identify additional B7-supertype epitopes, further studies were undertaken. The protein sequences from the fourteen HCV isolates utilized above were again examined to identify conserved, motif-containing 8- and 11-mers. The isolates were also examined for 9- and 10-mer sequences allowing for lower conservancy (51%-78%). Twenty-five 8-mers, sixteen 11-mers, and thirty-five 9- and 10-mers were identified, synthesized, and tested for binding to B*0702. Thirteen peptides bound with high or intermediate affinity ($IC_{50} \leq 500$ nM) (Table XXVIIIb). These peptides were additionally screened for binding to other B7-supertype molecules. Only one cross-reactive binder, the NS3 1378 8-mer (peptide 29.0035/1260.04), was identified (Table XXVIIIb).

In summary, a total of two cross-reactive B7-supertype binders were identified (Core 169 and NS3 1378).

Selection of A1 and A24 motif-bearing epitopes

To further increase population coverage, HLA-A1 and -A24 epitopes can also be incorporated into potential vaccine constructs.

In a previous analysis, two A1 and three A24 binders, 100% conserved among four strains of HCV, were identified (Wentworth *et al.*, *Int. Immunol.* 8:651-659, 1996). An analysis of the protein sequence data from the fourteen HCV strains utilized above demonstrated that these peptides were >79% conserved, and also identified an additional

- eleven A1- and twenty five A24-motif-containing conserved sequences (see Table XXIXA and B). Eight of the additional eleven A1 peptides and seven of the additional twenty five A24 peptides were tested for binding to the appropriate HLA molecule (*i.e.*, A1 or A24). Overall, as shown in Table XXIX, four A1-motif peptides (A) and three A24-motif peptides (B) have been found with binding capacities of 500 nM or less for the appropriate allele-specific HLA molecule.

- Analysis of the HLA-A2 and A3 supermotif-bearing epitopes identified above revealed that in 13/14 cases, peptides binding the supertype prototype HLA molecule (*i.e.* A*0201 for the A2 supertype, and A*0301 for the A3 supertype) with an IC_{50} of less than 100nM were cross-reactive and recognized by HCV-infected patients as described in Example 3, which follows. Based on these observations, two A1 peptides and one A24 peptide epitopes were also selected as candidates for inclusion in vaccine compositions; these peptides bind the appropriate HLA molecule with an IC_{50} of less than 100nM.

15 Example 3: Confirmation of Immunogenicity
Evaluation of A*0201 immunogenicity

It has been shown that CTL induced in A*0201/K^b transgenic mice exhibit specificity similar to CTL induced in the human system (*see, e.g.*, Vitiello *et al.*, *J. Exp. Med.* 173:1007-1015, 1991; Wentworth *et al.*, *Eur. J. Immunol.* 26:97-101, 1996).

- 20 Accordingly, these mice were used to evaluate the immunogenicity of the twelve conserved A2-supertype cross-reactive peptides identified in Example 2 above.

CTL induction in transgenic mice following peptide immunization has been described (Vitiello *et al.*, *J. Exp. Med.* 173:1007-1015, 1991; Alexander *et al.*; *J. Immunol.* 159:4753-4761, 1997). In these studies, mice were injected subcutaneously at the base of the tail with each peptide (50 µg/mouse) emulsified in IFA in the presence of an excess of an IA^b-restricted helper peptide (140 µg/mouse) (HBV core 128-140, Sette *et al.*, *J. Immunol.* 153:5586-5592, 1994). Eleven days after injection, splenocytes were incubated in the presence of peptide-loaded syngenic LPS blasts. After six days, cultures were assayed for cytotoxic activity using peptide-pulsed targets. The data, summarized in Table XXX, indicate that 7 of the 12 peptides (58%) were capable of inducing primary CTL responses in A*0201/K^b transgenic mice. (For these studies, a peptide was considered positive if it induced CTL (L.U. 30/10⁶ cells ≥2 in at least two transgenic animals (Wentworth *et al.*, *Eur. J. Immunol.* 26:97-101, 1996).

The conserved, cross reactive candidate CTL epitopes were also tested for recognition *in vitro* by PBMCs obtained from HCV-infected patients. Briefly, PBMC from patients infected with HCV were cultured in the presence of 10 µg/ml of synthetic peptide. After 7 and 14 days, the cultures were restimulated with peptide. The cultures were assayed for cytolytic activity on day 21 using target cells pulsed with the specific peptide in a standard four hour ⁵¹Cr release assay. The data are summarized in Table XXX. As shown, all 12 peptides are CTL epitopes recognized by PBMC from HCV-infected patients. From the data in Table XXX, it is interesting to note that HLA transgenics did not fully reveal the immunogenicity of some peptides that were positive in recall responses. This apparent discrepancy may reflect differences in the route of immunization utilized (e.g., natural infection versus peptide immunization), or CTL repertoire.

*Evaluation of A*03/A11 immunogenicity*

The immunogenicity of six of the eight A3-supertype cross-reactive peptides identified in Example 2 above was evaluated in HLA-A11/K^b transgenic mice, using the protocol described above for HLA-A2 transgenic mice (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). Five of these six peptides were able to induce primary CTL responses (Table XXXI).

All eight peptides were also studied by collaborators using PBMC cultures from HCV infected patients and contacts of such patients. This data is also summarized in Table XXXI. Briefly, all eight peptides were recognized by HCV infected individuals.

Evaluation of B7 immunogenicity

One of the two B7-supertype cross-reactive peptides (1145.12, Core 169) has been evaluated for immunogenicity in HCV-infected patients. Two independent collaborators have shown that this peptide is indeed immunogenic, and is recognized by T cells from HCV-infected patients (Chang *et al.*, *J. Immunol.* 162:1156-1164, 1999)

30 Example 4: Implementation of the Extended Supermotif to Improve the Binding Capacity of Native Epitopes by Creating Analogs

HLA motifs and supermotifs (comprising primary and/or secondary residues) are useful in the identification and preparation of highly cross-reactive native peptides, as demonstrated herein. Moreover, the definition of HLA motifs and supermotifs also

allows one to engineer highly cross-reactive epitopes by identifying residues within a native peptide sequence which can be analogued, or "fixed" to confer upon the peptide certain characteristics, *e.g.* greater cross-reactivity within the group of HLA molecules that comprise a supertype, and/or greater binding affinity for some or all of those HLA molecules. Examples of analog peptides that exhibit modulated binding affinity are set forth in this example.

Analoging at Primary Anchor Residues

As shown in Example 2, more than ten different HCV-derived, A2-supertype-restricted epitopes were identified. Peptide engineering strategies are implemented to further increase the cross-reactivity of the candidate epitopes identified above which bind 3/5 of the A2 supertype alleles tested. On the basis of the data disclosed, *e.g.*, in related and co-pending U.S.S.N 09/226,775, the main anchors of A2-supermotif-bearing peptides are altered, for example, to introduce a preferred L, I, V, or M at position 2, and I or V at the C-terminus.

To analyze the cross-reactivity of the analog peptides, each engineered analog is initially tested for binding to the prototype A2 supertype allele A*0201, then, if A*0201 binding capacity is maintained, for A2-supertype cross-reactivity.

Similarly, analogs of HLA-A3 supermotif-bearing epitopes may also be generated. For example, peptides binding to 3/5 of the A3-supertype molecules may be engineered at primary anchor residues to possess a preferred residue (V, S, M, or A) at position 2.

The analog peptides are then tested for the ability to bind A*03 and A*11 (prototype A3 supertype alleles). Those peptides that demonstrate ≤ 500 nM binding capacity are then tested for A3-supertype cross-reactivity.

Similarly to the A2- and A3- motif bearing peptides, peptides binding 3 or more B7-supertype alleles may be improved, where possible, to achieve increased cross-reactive binding. B7 supermotif-bearing peptides may, for example, be engineered to possess a preferred residue (V, I, L, or F) at the C-terminal primary anchor position, as demonstrated by Sidney *et al.* (*J. Immunol.* 157:3480-3490, 1996).

Analoging at Secondary Anchor Residues

Moreover, HLA supermotifs are of value in engineering highly cross-reactive peptides and/or peptides that bind HLA molecules with increased affinity by identifying

particular residues at secondary anchor positions that are associated with such properties. Demonstrating this, the binding capacity of a peptide representing a discreet single amino acid substitution at position one was analyzed. Peptide 1145.13 (Table XXVIIIc), which represents the substitution of L to F at position 1 of the core 169 sequence, binds all five B7-supertype molecules with a good affinity (all IC_{50} values ≤ 132 nM), and in 3 instances has higher affinity over that of the parent peptide by >35 -fold.

Because so few B7-supertype cross-reactive epitopes were identified, our results from previous binding evaluations were analyzed to identify conserved (8-, 9-, 10-, or 11-mer) peptides which bind, minimally, 3/5 B7 supertype molecules with weak affinity (IC_{50} of 500nM-5 μ M). This analysis identified 9 peptides, 6 of which are analogues (including core 169 which had been previously analogued). These peptides are tested for enhanced binding affinity and B7-supertype cross-reactivity.

Engineered analogs with sufficiently improved binding capacity or cross-reactivity are tested for immunogenicity in HLA-B7-transgenic mice, following for example, IFA immunization or lipopeptide immunization.

In conclusion, these data demonstrate that by the use of even single amino acid substitutions, it is possible to increase the binding affinity and/or cross-reactivity of peptide ligands for HLA supertype molecules.

20 Example 5: Identification of conserved HCV-derived sequences with HLA-DR binding motifs

Peptide epitopes bearing an HLA class II supermotif or motif may also be identified as outlined below using methodology similar to that described in Examples 1-3.

25 *Selection of HLA-DR-supermotif-bearing epitopes*

To identify HCV-derived, HLA class II HTL epitopes, the same fourteen HCV polyprotein sequences used for the identification of HLA Class I supermotif/motif sequences were analyzed for the presence of sequences bearing an HLA-DR-motif or supermotif. Specifically, 15-mer sequences were selected comprising a DR-supermotif, further comprising a 9-mer core, and three-residue N- and C-terminal flanking regions (15 amino acids total). It was also required that the 15-mer sequence be conserved in at least 79% (11/14) of the HCV strains analyzed. These criteria identified a total of 49 non-redundant sequences, which are shown in Table XXXIIIA. (In the context of Class II

epitopes, a sequence is considered operationally redundant if more than 80% of it's sequence overlaps with another peptide.)

5 Protocols for predicting peptide binding to DR molecules have been developed (Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). These protocols, specific for individual DR molecules, allow the scoring, and ranking, of 9-mer core regions. Each protocol not only scores peptide sequences for the presence of DR-supermotif primary anchors (i.e., at position 1 and position 6) within a 9-mer core, but additionally evaluates sequences for the presence of secondary anchors. Using allele specific selection tables (see, e.g., Southwood *et al.*, *ibid.*), it has been found that these protocols efficiently select 10 peptide sequences with a high probability of binding a particular DR molecule. Additionally, it has been found that performing these protocols in tandem, specifically those for DR1, DR4w4, and DR7, can efficiently select DR cross-reactive peptides.

20 To see if these protocols serve to identify additional epitopes, the same HCV polyproteins used above were re-scanned for the presence of 15-mer peptides with 9-mer core regions that were $\geq 79\%$ (11/14 strains) conserved. This identified 152 sequences; 49 of which were identified previously, as described above. Next, the 9-mer core region of each of these peptides was scored using the DR1, DR4w4, and DR7 algorithms. Twenty-two peptides, including 12 new sequences (10 peptides were from the original set of 49) were found to have 9-mer cores with protocol-derived scores predictive of cross-reactive 20 DR binders. The 12 additional sequences are shown in Table XXXIIB.

25 The conserved, HCV-derived peptides identified above were tested for their binding capacity for various common HLA-DR molecules. All peptides were initially tested for binding to the DR molecules in the primary panel: DR1, DR4w4, and DR7. Peptides binding at least 2 of these 3 DR molecules were then tested for binding to DR2w2 β 1, DR2w2 β 2, DR6w19, and DR9 molecules in secondary assays. Finally, peptides binding at least 2 of the 4 secondary panel DR molecules, and thus cumulatively at least 4 of 7 different DR molecules, were screened for binding to DR4w15, DR5w11, and DR8w2 molecules in tertiary assays. Peptides binding at least 7 of the 10 DR molecules comprising the primary, secondary, and tertiary screening assays were 30 considered cross-reactive DR binders. The composition of these screening panels, and the phenotypic frequency of associated antigens, are shown in Table XXXIII.

Upon testing, it was found that 29 of the original 75 peptides (39%) bound two or more of the primary HLA molecules. Twenty-six of these cross-reactive binders were

then tested in the secondary assays, and nineteen were found to bind at least four of the seven HLA DR molecules in the primary and secondary panels. Finally, the nineteen peptides passing the secondary screening phase were tested for binding in the tertiary assays. As a result, nine peptides were identified which bound at least seven of ten common HLA-DR molecules. Table XXXIV shows these nine peptides and their binding capacity for each allele-specific HLA-DR molecule in the primary through tertiary panels. Also shown in Table XXXIV are two peptides (F134.05 and F134.08) for which a complete binding analysis was not performed. However, both of these peptides bound six of the seven HLA DR molecules tested. F134.08 nests peptide 1283.44, which bound eight of 10 allele-specific HLA molecules.

In conclusion, eleven cross-reactive DR-binding peptides, derived from six discrete (*i.e.* non-redundant) regions of the HCV genome, have been identified. Two of the six regions from which these epitopes were derived are covered by multiple, overlapping epitopes.

Selection of conserved DR3 motif peptides

Because HLA-DR3 is an allele that is prevalent in Caucasian, Black, and Hispanic populations, DR3 binding capacity is an important criterion in the selection of HTL epitopes. However, data generated previously indicated that DR3 only rarely cross-reacts with other DR alleles (Sidney *et al.*, *J. Immunol.* 149:2634-2640, 1992; Geluk *et al.*, *J. Immunol.* 152:5742-5748, 1994; Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). This is not entirely surprising in that the DR3 peptide-binding motif appears to be distinct from the specificity of most other DR alleles.

To efficiently identify peptides that bind DR3, target proteins were analyzed for conserved sequences carrying one of the two DR3 specific binding motifs reported by Geluk *et al.* (*J. Immunol.* 152:5742-5748, 1994). Fifteen sequences, including a peptide nested within a DR-supermotif sequence identified above (peptide Pape 22), were identified (Table XXXIIId). Preferably, DR3 motifs will be found clustered in proximity with DR supermotif regions.

Fourteen of the fifteen peptides containing a DR3 motif were tested for their DR3 binding capacity. Two peptides (CH35.0106 and CH35.0107) were found to bind DR3 with an affinity of $1\mu\text{M}$ or less (Table XXXV), and thereby qualify as HLA class II high affinity binders.

DR3 binding epitopes identified in this manner may then be included in vaccine compositions with DR supermotif-bearing peptide epitopes.

Example 6: Immunogenicity of candidate HCV-derived HTL epitopes and known dominant HCV HTL epitope

In the course of collaborative studies with G. Pape and C. Ferrari, eight conserved, HCV-derived peptides have been identified which are recognized by HCV-infected individuals.

One of these studies (Diepolder *et al.*, *J. Virol.* 71:6011-6019, 1997), identified peptide F98.05, which spans residues 1248-1261 of the NS3 protein, as an immunodominant CD4+ T-cell epitope that was recognized by 14/23 NS3-specific CD4+ T-cell clones from 4/5 patients with acute hepatitis C infection. This epitope, shown above to be an HLA-DR cross-reactive binder (see Table XXXIV), was capable of being presented to helper CD4+ T cells by multiple HLA molecules (DR4, DR11, DR12, DR13, and DR16). Two other peptides, Pape 22 and Pape 29, were also recognized by CD4+ T cell clones, although, in a more limited context; correspondingly, neither of these peptides are DR-cross-reactive binders.

By direct peripheral blood T cell stimulation and by fine specificity analysis of HCV-specific T-cell lines and clones, studies done in collaboration with Ferrari's group identified 6 immunodominant epitopes, including one also identified in the Pape collaboration, that are derived from conserved regions of the core, NS3, and NS4 proteins. These epitopes were also found to be cross-reactive, being presented to T cells in the context of different Class II molecules. Three of the 6 epitopes, F98.04 (F134.03), F134.05 and F134.08, are cross-reactive HLA-DR binders (see Table XXXIV).

In conclusion, the immunogenicity of 8 epitopes derived from conserved regions of the HCV genome has been demonstrated. Three of these epitopes (F98.05, F134.05, and F134.08; see Table XXXIV) are broadly cross-reactive HLA-DR binding peptides.

Example 7. Calculation of phenotypic frequencies of HLA-supertypes in various ethnic backgrounds to determine breadth of population coverage

This example illustrates the assessment of the breadth of population coverage of a vaccine composition comprised of multiple epitopes comprising multiple supermotifs and/or motifs.

In order to analyze population coverage, gene frequencies of HLA alleles were determined. Gene frequencies for each HLA allele were calculated from antigen or allele frequencies utilizing the binomial distribution formulae $gf=1-(\text{SQRT}(1-af))$ (see, e.g., Sidney *et al.*, *Human Immunol.* 45:79-93, 1996). To obtain overall phenotypic frequencies, cumulative gene frequencies were calculated, and the cumulative antigen frequencies derived by the use of the inverse formula $[af=1-(1-Cgf)^2]$.

- Where frequency data was not available at the level of DNA typing, correspondence to the serologically defined antigen frequencies was assumed. To obtain total potential supertype population coverage no linkage disequilibrium was assumed, and only alleles confirmed to belong to each of the superotypes were included (minimal estimates). Estimates of total potential coverage achieved by inter-loci combinations were made by adding to the A coverage the proportion of the non-A covered population that could be expected to be covered by the B alleles considered (e.g., $\text{total}=A+B*(1-A)$). Confirmed members of the A3-like supertype are A3, A11, A31, A*3301, and A*6801. Although the A3-like supertype may also include A34, A66, and A*7401, these alleles were not included in overall frequency calculations. Likewise, confirmed members of the A2-like supertype family are A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802, and A*6901. Finally, the B7-like supertype-confirmed alleles are: B7, B*3501-03, B51, B*5301, B*5401, B*5501-2, B*5601, B*6701, and B*7801 (potentially also B*1401, B*3504-06, B*4201, and B*5602).

- Population coverage achieved by combining the A2-, A3- and B7-supertypes is approximately 86% in five major ethnic groups (see Table XXI). Coverage may be extended by including peptides bearing the A1 and A24 motifs. On average, A1 is present in 12% and A24 in 29% of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Together, these alleles are represented with an average frequency of 39% in these same ethnic populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95%. An analogous approach can be used to estimate population coverage achieved with combinations of class II motif-bearing epitopes.

Summary of candidate HLA class I and class II epitopes

In summary, on the basis of the data presented in the above examples, 26 CTL candidate peptide epitopes derived from conserved regions of the HCV virus have been

identified (Table XXXVIa). These include twelve HLA-A2 supermotif-bearing epitopes, eight HLA-A3 supermotif-bearing epitopes, and one HLA-B7 supermotif-bearing epitope, each capable of binding to multiple A2-, A3-, or B7-supertype molecules, and immunogenic in HLA transgenic mice or antigenic for human PBL (with the exception of peptide 29.0035/1260.04). Additional epitopes not evaluated for immunogenicity are also included. They are an additional B7-supermotif-bearing epitope and two HLA-A1 and one HLA-A24 high-affinity binding peptides. A known HLA-A31 restricted epitope (VGIYLLPNR), which also binds HLA-A33, is also set out in Table XXXVIa and is useful in combination with other Class I or Class II epitopes.

With these 26 CTL epitopes (as disclosed herein and from the art), average population coverage, (*i.e.*, recognition of at least one HCV epitope), is predicted to be greater than 95% in each of five major ethnic populations. The potential redundancy of coverage afforded by 25 of these epitopes (the peptide 24.0086 was not included) was estimated using the game theory Monte Carlo simulation analysis, which is known in the art (see *e.g.*, Osborne, M.J. and Rubinstein, A. "A course in game theory" MIT Press, 1994). As shown in Figure 1, it is estimated that 90% of the individuals in a population comprised of the Caucasian, North American Black, Japanese, Chinese, and Hispanic ethnic groups would recognize 2 or more of the candidate epitopes described herein.

A list of HCV-derived HTL epitopes that would be preferred for use in the design of minigene constructs or other vaccine formulations is summarized in Table XXXVIb. As shown, 9 different peptide-binding regions have been identified which bind multiple HLA-DR molecules or bind HLA-DR3. (In the case of the NS4 1914-1935 region, the longer peptide, F134.08, recognized by patients, was chosen over the shorter peptide, 1283.44. The longer peptide essentially incorporates the shorter peptide, and also binds additional DR molecules that the shorter peptide does not bind.) Three of these peptides have been recognized as dominant epitopes in HCV infected patients.

It is estimated that each of 10 common DR molecules recognizing the DR supermotif, and DR3, are covered by a minimum of 2 epitopes. Correspondingly, the total estimated population coverage represented by this panel of epitopes is in excess of 91% in each of the 5 major ethnic populations (Table XXXVII).

Example 8: Recognition Of Generation Of Endogenous Processed Antigens After Priming

This example determines that CTL induced by native or analogued peptide epitopes identified and selected as described in Examples 1-6 recognize endogenously synthesized, *i.e.*, native antigens.

Effector cells isolated from transgenic mice that are immunized with peptide epitopes as in Example 3, for example HLA-A2 supermotif-bearing epitopes, are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ⁵¹Cr labeled Jurkat-A2.1/K^b target cells in the absence or presence of peptide, and also tested on ⁵¹Cr labeled target cells bearing the endogenously synthesized antigen, *i.e.* cells that are stably transfected with HCV expression vectors.

The result will demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized HCV antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that is being evaluated. In addition to HLA-A*0201/K^b transgenic mice, several other transgenic mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

Example 9: Activity Of CTL-HTL Conjugated Epitopes In Transgenic Mice

This example illustrates the induction of CTLs and HTLs in transgenic mice by use of an HCV CTL/HTL peptide conjugate whereby the vaccine composition comprises peptides administered to an HCV-infected patient or an individual at risk for HCV. The peptide composition can comprise multiple CTL and/or HTL epitopes. This analysis demonstrates enhanced immunogenicity that can be achieved by inclusion of one or more HTL epitopes in a vaccine composition. Such a peptide composition can comprise a lipidated HTL epitope conjugated to a preferred CTL epitope containing, for example, at least one CTL epitope selected from Table XXVI-XXIX, or an analog of that epitope. The HTL epitope is, for example, selected from Table XXXII.

Lipopeptide preparation: Lipopeptides are prepared by coupling the appropriate fatty acid to the amino terminus of the resin bound peptide. A typical procedure is as

follows: A dichloromethane solution of a four-fold excess of a pre-formed symmetrical anhydride of the appropriate fatty acid is added to the resin and the mixture is allowed to react for two hours. The resin is washed with dichloromethane and dried. The resin is then treated with trifluoroacetic acid in the presence of appropriate scavengers [e.g. 5% (v/v) water] for 60 minutes at 20°C. After evaporation of excess trifluoroacetic acid, the crude peptide is washed with diethyl ether, dissolved in methanol and precipitated by the addition of water. The peptide is collected by filtration and dried.

- Immunization procedures: Immunization of transgenic mice is performed as described (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are useful for the assessment of the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, are primed subcutaneously (base of the tail) with 0.1 ml of peptide conjugate formulated in saline, or DMSO/saline. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated lymphoblasts coated with peptide.

Cell lines: Target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (*e.g.*, Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991)

- In vitro* CTL activation: One week after priming, spleen cells (30x10⁶ cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated lymphoblasts (10x10⁶ cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

- Assay for cytotoxic activity: Target cells (1.0 to 1.5x10⁶) are incubated at 37°C in the presence of 200 µl of ⁵¹Cr. After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of 1 µg/ml. For the assay, 10⁴ ⁵¹Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200 µl) in U-bottom 96-well plates. After a 6 hour incubation period at 37°C, a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release = 100 x (experimental release - spontaneous release)/(maximum release - spontaneous release). To facilitate comparison between separate CTL assays run under the same conditions, % ⁵¹Cr release data is expressed as lytic units/10⁶ cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a 6

hour ^{51}Cr release assay. To obtain specific lytic units/ 10^6 , the lytic units/ 10^6 obtained in the absence of peptide is subtracted from the lytic units/ 10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., 5×10^5 effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e., 5×10^4 effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: $[(1/50,000) - (1/500,000)] \times 10^6 = 18 \text{ LU}$.

The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using the CTL epitope as outlined in Example 3. Analyses similar to this may be performed to evaluate the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

Example 10. Selection of CTL and HTL epitopes for inclusion in an HCV-specific vaccine.

This example illustrates the procedure for the selection of peptide epitopes for vaccine compositions of the invention. The peptides in the composition can be in the form of a nucleic acid sequence, either single or one or more sequences (i.e., minigene) that encodes peptide(s), or may be single and/or polypeptidic peptides.

Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with tumor clearance. For example, vaccine can include 3-4 epitopes that come from at least one HCV antigen region. Epitopes from one region can be used in combination with epitopes from one or more additional HCV antigen regions. Analogs of epitopes can also be selected for inclusion in the vaccine.

Epitopes are often selected that have a binding affinity of an IC_{50} of 500 nM or less for an HLA class I molecule, or for class II, an IC_{50} of 1000 nM or less.

Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess breadth, or redundancy, of population coverage.

- When creating a polyepitopic compositions, *e.g.* a minigene, it is typically desirable to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same, as those employed when selecting a peptide comprising nested epitopes. Additionally, however, upon
- 5 determination of the nucleic acid sequence to be provided as a minigene, the peptide sequence encoded thereby is analyzed to determine whether any "junctional epitopes" have been created. A junctional epitope is a potential HLA binding epitope, as predicted, *e.g.*, by motif analysis. Junctional epitopes are generally to be avoided because the recipient may bind to an HLA molecule and generate an immune response to that epitope,
- 10 which is not present in a native protein sequence.

- Peptide epitopes for inclusion in vaccine compositions are, for example, selected from those listed in Tables XXVI-XXIX and Table XXXII. A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude of an immune response that clears an acute HCV
- 15 infection.

Example 11: Construction of Minigene Multi-Epitope DNA Plasmids

- This example provides guidance for the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of CTL
- 20 and/or HTL epitopes or epitope analogs as described herein. Examples of the construction and evaluation of expression plasmids are described, for example, in pending U.S.S.N. 09/311,784 filed 5/13/99. An example of such a plasmid for the expression of HCV epitopes is shown in Figure 2, which illustrates the orientation of HCV peptide epitopes in a minigene construct.

- 25 A minigene expression plasmid may include multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes (Figure 2). Preferred epitopes are identified, for example, in Tables XXVI-XXIX and XXXII. HLA class I supermotif or
- 30 motif-bearing peptide epitopes derived from multiple HCV antigens, *e.g.*, the core, NS4, NS3, NS5, NS1/E2, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from multiple HCV antigens to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for

inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

This example illustrates the methods to be used for construction of such a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

The minigene DNA plasmid contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the pcDNA 3.1 Myc-His vector.

Overlapping oligonucleotides, for example eight oligonucleotides, averaging approximately 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated T_m of each primer pair) for 30 sec, and 72°C for 1 min.

For the first PCR reaction, 5 µg of each of two oligonucleotides, *i.e.*, an amplification primer pair, are annealed and extended: Oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 µl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM (NH₄)₂SO₄, 20 mM Tris-chloride, pH 8.75, 2 mM MgSO₄, 0.1% Triton X-100, 100 µg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product for 25 additional cycles. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

Example 12. The plasmid construct and the degree to which it induces immunogenicity.

The degree to which the plasmid construct prepared using the methodology outlined in Example 11 is able to induce immunogenicity is evaluated through *in vivo*

injections into mice and subsequent *in vitro* assessment of CTL and HTL activity, which are analysed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.*, in U.S.S.N. 09/311,784 filed 5/13/99 and Alexander *et al.*, *Immunity* 1:751-761, 1994. For example, to assess the capacity of a pMin minigene construct that contains HLA-A2

- 5 supermotif epitopes to induce CTLs *in vivo*, HLA-A2.1/K^b transgenic mice are immunized intramuscularly with 100 µg of naked cDNA. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.
- 10 Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polypeptidic peptide), then assayed for peptide-specific cytotoxic activity in a ⁵¹Cr release assay. The results indicate the magnitude of the CTL response directed against the A3-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and
- 15 polypeptidic vaccine. It is, therefore, found that the minigene elicits immune responses directed toward the HLA-A2 supermotif peptide epitopes as does the polypeptidic peptide vaccine. A similar analysis is also performed using other HLA-A3 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes.

- 20 To assess the capacity of a class II epitope encoding minigene to induce HTLs *in vivo*, I-A^b restricted mice, for example, are immunized intramuscularly with 100 µg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant.

- 25 CD4+ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ³H-thymidine incorporation proliferation assay, (*see, e.g.*, Alexander *et al.* *Immunity* 1:751-761, 1994). the results indicate the magnitude of the HTL response, thus demonstrating the *in vivo* immunogenicity of the minigene.

- 30 Alternatively, plasmid constructs can be evaluated *in vitro* by testing for epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines "antigenicity" and allows the use of human APC. The assay determines the ability of the epitope to be presented by the

APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (*see, e.g., Sijts et al., J. Immunol.* 156:683-692, 1996; Demotz *et al., Nature* 342:682-684, 1989); or the number of peptide-HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by infected or transfected target cells, and then determining the concentration of peptide necessary to obtained equivalent levels of lysis or lymphokine release (*see, e.g., Kageyama et al., J. Immunol.* 154:567-576, 1995).

10 Example 13: Peptide Composition for Prophylactic Uses

Vaccine compositions of the present invention are used to prevent HCV infection in persons who are at risk for such infection. For example, a polypeptidic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes such as those selected in Examples 9 and/or 10, which are also selected to target greater than 80% of the population, is administered to individuals at risk for HCV infection. The composition is provided as a single lipidated polypeptide that encompasses multiple epitopes. The vaccine is administered in an aqueous carrier comprised of Freund's Incomplete Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 µg, generally 100-5,000 µg, for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against HCV infection.

Alternatively, the polypeptidic peptide composition can be administered as a nucleic acid in accordance with methodologies known in the art and disclosed herein.

Example 14: Polypeptidic Vaccine Compositions Derived from Native HCV Sequences

A native HCV polyprotein sequence is screened, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify "relatively short" regions of the polyprotein that comprise multiple epitopes and is preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct, even overlapping, epitopes is selected and used to generate a minigene construct. The construct is engineered to express the peptide, which

corresponds to the native protein sequence. The "relatively short" peptide is generally less than 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has

5 maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping (*i.e.*, frame shifted relative to one another). For example, with frame shifted overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic

10 purposes.

The vaccine composition will preferably include, for example, three CTL epitopes and at least one HTL epitope from an HCV antigen. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the

15 epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune

20 response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent analogs) directs the immune response to multiple peptide sequences that are actually present in native HCV antigens thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of

25 scale when producing nucleic acid vaccine compositions.

Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

30 Example 15. Polyepitopic Vaccine Compositions Directed To Multiple Diseases

The HCV peptide epitopes of the present invention are used in conjunction with peptide epitopes from target antigens related to one or more other diseases, to create a vaccine composition that is useful for the prevention or treatment of HCV as well as the

one or more other disease(s). Examples of the other diseases include, but are not limited to, HIV, and HBV.

- For example, a polypeptidic peptide composition comprising multiple CTL and HTL epitopes that target greater than 98% of the population may be created for administration to individuals at risk for both HCV and HIV infection. The composition can be provided as a single polypeptide that incorporates the multiple epitopes from the various disease-associated sources, or can be administered as a composition comprising one or more discrete epitopes.

10 Example 16. Use of peptides to evaluate an immune response

- Peptides of the invention may be used to analyze an immune response for the presence of specific CTL or HTL populations directed to a prostate cancer-associated antigen. Such an analysis may be performed using multimeric complexes as described, e.g., by Ogg *et al.*, *Science* 279:2103-2106, 1998 and Greten *et al.*, *Proc. Natl. Acad. Sci. USA* 95:7568-7573, 1998. In the following example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

- In this example, highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, HCV HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of disease or following immunization using an HCV peptide containing an A*0201 motif. Tetrameric complexes are synthesized as described (Musey *et al.*, *N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β 2-microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β 2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5'triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

- For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300g for 5 minutes and resuspended in 50 μ l of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive uninfected donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the HCV epitope, and thus the stage of HCV infection or exposure to a vaccine that elicits a protective or therapeutic response.

Example 17: Use of Peptide Epitopes to Evaluate Recall Responses

- The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from infection, who are chronically infected with HCV, or who have been vaccinated with an HCV vaccine.

- For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any HCV vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that are preferably highly conserved and, optimally, bear supermotifs to provide cross-reactivity with multiple HLA supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

- PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 μ g/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 μ g/ml to each well and HBV core 128-140 epitope is added at 1 μ g/ml to each well as a source of T cell help during the first week of stimulation.

In the microculture format, 4×10^5 PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 μ l/well of complete RPMI. On

days 3 and 10, 100 ml of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with uninfected control subjects as previously described (Rehermann, *et al.*, *Nature Med.* 2:1104,1108, 1996; Rehermann *et al.*, *J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al.* *J. Clin. Invest.* 98:1432-1440, 1996).

Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al.* *J. Virol.* 66:2670-2678, 1992).

Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 μM , and labeled with 100 μCi of ^{51}Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS.

Cytolytic activity is determined in a standard 4-h, split well ^{51}Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: $100 \times [(\text{experimental release} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})]$. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is $<25\%$ of maximum release for all experiments.

The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to HCV or an HCV vaccine.

The class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of 1.5×10^5 cells/well and are stimulated with 10 $\mu\text{g/ml}$ synthetic peptide, whole antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1 μCi ^3H -thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^3H -thymidine

incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^3H -thymidine incorporation in the presence of antigen divided by the ^3H -thymidine incorporation in the absence of antigen.

5 Example 18: Induction Of Specific CTL Response In Humans

A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

- 10 A total of about 27 subjects are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

- 15 Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

- 20 The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

- 25 Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

- 30 The vaccine is found to be both safe and efficacious.

Example 19: Phase II Trials In Patients Infected With HCV

Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to patients having chronic HCV infection. The main objectives of

the trials are to determine an effective dose and regimen for inducing CTLs in chronically infected HCV patients, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of chronically infected CTL patients, as manifested by a transient flare in alanine aminotransferase (ALT), normalization of ALT, and reduction in HCV DNA. Such a study is designed, for example, as follows:

The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65, include both males and females, and represent diverse ethnic backgrounds. All of them are infected with HCV for over five years and are HIV, HBV and delta hepatitis virus (HDV) negative, but have positive levels of HCV antigen.

The magnitude and incidence of ALT flares and the levels of HCV DNA in the blood are monitored to assess the effects of administering the peptide compositions. The levels of HCV DNA in the blood are an indirect indication of the progress of treatment. The vaccine composition is found to be both safe and efficacious in the treatment of chronic HCV infection.

Example 20. Induction of CTL Responses Using a Prime Boost Protocol

A prime boost protocol can also be used for the administration of the vaccine to humans. Such a vaccine regimen may include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

For example, the initial immunization may be performed using an expression vector, such as that constructed in Example 11, in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μ g) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is administered. The booster can, e.g., be recombinant fowlpox virus administered at a dose of $5 \cdot 10^7$ to $5 \cdot 10^9$ pfu. An alternative

- recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polypeptidic protein or a mixture of the peptides can be administered. For evaluation of vaccine efficacy, patient blood samples will be obtained before immunization as well as at intervals following administration of the
- 5 initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.
- Analysis of the results will indicate that a magnitude of response sufficient to
- 10 achieve protective immunity or to treat HCV infection is generated.

Example 21. Administration of Vaccine Compositions Using Dendritic Cells

- Vaccines comprising peptide epitopes of the invention may be administered using dendritic cells. In this example, the peptide-pulsed dendritic cells can be administered to
- 15 a patient to stimulate a CTL response *in vivo*. In this method dendritic cells are isolated, expanded, and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses *in vivo*. The induced CTL and HTL then destroy (CTL) or facilitate destruction (HTL) of the specific target HCV-infected cells that bear the proteins from which the
- 20 epitopes in the vaccine are derived.

- Alternatively, *Ex vivo* CTL or HTL responses to a particular tumor-associated antigen can be induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells, such as dendritic cells, and the appropriate immunogenic peptides. After an
- 25 appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cells, *i.e.*, tumor cells.

30 Example 22: Alternative Method of Identifying Motif-Bearing Peptides

Another way of identifying motif-bearing peptides is to elute them from cells bearing defined MHC molecules. For example, EBV transformed B cell lines used for tissue typing, have been extensively characterized to determine which HLA molecules they express. In certain cases these cells express only a single type of HLA molecule.

These cells can then be infected with a pathogenic organism, *e.g.*, HCV, or transfected with nucleic acids that express the antigen of interest. Thereafter, peptides produced by endogenous antigen processing of peptides produced consequent to infection (or as a result of transfection) will bind be displayed on the cell surface. The peptides are then eluted from the HLA molecules by exposure to mild acid conditions and their amino acid sequence determined, *e.g.*, by mass spectral analysis (*e.g.*, Kubo *et al.*, *J. Immunol.* 152:3913, 1994). Because, as disclosed herein, the majority of peptides that bind a particular HLA molecule are motif-bearing, this is an alternative modality for obtaining the motif-bearing peptides correlated with the particular HLA molecule expressed on the cell.

Alternatively, cell lines that do not express any endogenous HLA molecules can be transfected with an expression construct encoding a single HLA allele. These cells may then be used as described, *i.e.*, they may be infected with a pathogenic organism or transfected with nucleic acid encoding an antigen of interest to isolate peptides corresponding to the pathogen or antigen of interest that have been presented on the cell surface. Peptides obtained from such an analysis will bear motif(s) that correspond to binding to the single HLA allele that is expressed in the cell.

As appreciated by one in the art, one can perform a similar analysis on a cell bearing more than one HLA allele and subsequently determine peptides specific for each HLA allele expressed. Moreover, one of skill would also recognize that means other than infection or transfection, such as loading with a protein antigen, can be used to provide a source of antigen to the cell.

The above examples are provided to illustrate the invention but not to limit its scope. For example, the human terminology for the Major Histocompatibility Complex, namely HLA, is used throughout this document. It is to be appreciated that these principles can be extended to other species as well. Thus, other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent application cited herein are hereby incorporated by reference for all purposes.

TABLE I

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	T , I , L , <i>V</i> , M , S		F , W , Y
A2	L , I , V , M , A , T , <i>Q</i>		I , V , M , A , T , L
A3	V , S , M , A , T , L , <i>I</i>		R , K
A24	Y , F , <i>W</i> , <i>I</i> , <i>V</i> , L , M , T		F , I , Y , <i>W</i> , L , M
B7	P		V , I , L , F , M , <i>W</i> , Y , A
B27	R , H , K		F , Y , L , <i>W</i> , M , I , <i>V</i> , A
B44	E , D		F , W , L , I , M , V , A
B58	A , T , S		F , W , Y , L , <i>I</i> , <i>V</i> , M , A
B62	Q , L , <i>I</i> , <i>V</i> , M , P		F , W , Y , M , I , <i>V</i> , L , A
MOTIFS			
A1	T , S , M		Y
A1		D , E , A , S	Y
A2.1	L , M , <i>V</i> , <i>Q</i> , I , A , T		V , L , I , M , A , T
A3	L , M , V , I , S , A , T , F , C , G , D		K , Y , R , H , F , A
A11	V , T , M , L , I , S , A , G , N , C , D , F		K , R , Y , H
A24	Y , F , W , M		F , L , I , W
A*3101	M , V , T , A , L , I , S		R , K
A*3301	M , V , A , L , F , I , S , T		R , K
A*6801	A , V , T , M , S , L , I		R , K
B*0702	P		L , M , F , <i>W</i> , Y , A , I , <i>V</i>
B*3501	P		L , M , F , W , Y , I , <i>V</i> , A
B51	P		L , I , V , F , <i>W</i> , Y , A , M
B*5301	P		I , M , F , W , Y , A , L , <i>V</i>
B*5401	P		A , T , I , V , L , M , F , W , Y

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE II

	POSITION							
	1	2	3	4	5	6	7	8
SUPERMOTIFS								
A1		1° Anchor T,I,L,V,M,S						1° Anchor F,W,Y
A2		1° Anchor L,I,V,M,A						1° Anchor L,I,V,M,A,T
A3	preferred	1° Anchor V,S,M,A,T, L,I	Y,F,W (4/5)			Y,F,W (3/5)	Y,F,W (4/5)	1° Anchor R,K
	deleterious	D,E (3/5); P (5/5)		D,E (4/5)				
A24		1° Anchor Y,F,W,I,V, L,M,T						1° Anchor F,I,Y,W,L,M
B7	preferred	F,W,Y (5/5); L,I,V,M (3/5)	1° Anchor P	F,W,Y (4/5)			F,W,Y (3/5)	1° Anchor V,I,L,L,F,M,W,Y,A
	deleterious	D,E (3/5); P (5/5); G (4/5); A (3/5); Q,N (3/5)		D,E (3/5)	G (4/5)		Q,N (4/5)	D,E (4/5)
B27		1° Anchor R,H,K						1° Anchor F,Y,L,W,M,V,A
B44		1° Anchor E,D						1° Anchor F,W,Y,L,I,M,V,A
B58		1° Anchor A,T,S						1° Anchor F,W,Y,L,I,V,M,A
B62		1° Anchor Q,L,I,V,M, P						1° Anchor F,W,Y,M,I,V,L,A

POSITION

	1	2	3	4	5	6	7	8	C-terminus
MOTES									
A1 preferred 9-mer	G,F,Y,W	I ^o Anchor S,T,M	D,E,A	Y,F,W		P	D,E,Q,N	Y,F,W	I ^o Anchor Y
deleterious	D,E		R,H,K,L,I,V M,P	A	G	A			
A1 preferred 9-mer	G,R,H,K	A,S,T,C,L,I V,M	I ^o Anchor D,E,A,S	G,S,T,C		A,S,T,C	L,I,V,M	D,E	I ^o Anchor Y
deleterious	A	R,H,K,D,E, F,Y,F,W		D,E	P,Q,N	R,H,K	P,G	G,P	

POSITION

	1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A1 10-mer	preferred Y,F,W	1°Anchor S,T,M	D,E,A,Q,N	A	Y,F,W,Q,N	P,A,S,T,C	G,D,E	P	1°Anchor Y	
	deleterious G,P	R,H,K,G,L,I V,M	D,E	R,H,K	Q,N,A	R,H,K,Y,F, W	R,H,K	A		
A1 10-mer	preferred Y,F,W	S,T,C,L,I,V M	1°Anchor D,E,A,S	A	Y,F,W	P,G	G	Y,F,W	1°Anchor Y	
	deleterious R,H,K	R,H,K,D,E, P,Y,F,W		P	G	P,R,H,K	Q,N			
A2.1 9-mer	preferred Y,F,W	1°Anchor L,M,I/V,Q, A,T	Y,F,W	S,T,C	Y,F,W	A	P	1°Anchor V,L,I,M,A,T		
	deleterious D,E,P		D,E,R,K,H		R,K,H	D,E,R,K,H				
A2.1 10-mer	preferred A,Y,F,W	1°Anchor L,M,I/V,Q, A,T	L,V,I,M	G	G		F,Y,W, L,V,I,M	1°Anchor V,L,I,M,A,T		
	deleterious D,E,P		D,E	R,K,H,A	P	R,K,H	D,E,R, K,H	R,K,H		

POSITION									
	1	2	3	4	5	6	7	8	9 or C-terminus I ^o Anchor K,Y,R,H,F,A
A3 preferred	R,H,K	I ^o Anchor L,M,V,I,S, A,T,F,C,G D	Y,F,W	P,R,H,K,Y, F,W	A	Y,F,W		P	
deleterious	D,E,P		D,E						
A11 preferred	A	I ^o Anchor V,T,L,M,I, S,A,G,N,C, D,F	Y,F,W	Y,F,W	A	Y,F,W	Y,F,W	P	I ^o Anchor K _n RYH
deleterious	D,E,P						A	G	
A24 preferred 9-mer	Y,F,W,R,H,K	I ^o Anchor Y,F,W,M		S,T,C			Y,F,W	Y,F,W	I ^o Anchor F,L,I,W
deleterious	D,E,G		D,E	G	Q,N,P	D,E,R,H,K	G	A,Q,N	
A24 preferred 10-mer		I ^o Anchor Y,F,W,M		P	Y,F,W,P		P		I ^o Anchor F,L,I,W
deleterious			G,D,E	Q,N	R,H,K	D,E	A	Q,N	D,E,A
A3101 preferred	R,H,K	I ^o Anchor M,V,T,A,L, I,S	Y,F,W	P		Y,F,W	Y,F,W	A,P	I ^o Anchor R,K
deleterious	D,E,P		D,E		A,D,E	D,E	D,E	D,E	

POSITION									
	1	2	3	4	5	6	7	8	9 or C-terminus
A3301 preferred		1°Anchor M,V,A,L,F, I,S,T	Y,F,W				A,Y,F,W		1°Anchor R,K
deleterious	G,P		D,E						
A6801 preferred	Y,F,W,S,T,C	1°Anchor A,V,T,M,S, L,I			Y,F,W,L,I, V,M		Y,F,W	P	1°Anchor R,K
deleterious	G,P		D,E,G		R,H,K			A	
B0702 preferred	R,H,K,F,W,Y	1°Anchor P	R,H,K		R,H,K	R,H,K	R,H,K	P,A	1°Anchor L,M,F,W,Y,A, I,Y
deleterious	D,E,Q,N,P		D,E,P	D,E	D,E	G,D,E	Q,N	D,E	
B3501 preferred	F,W,Y,L,I,V,M	1°Anchor P	F,W,Y				F,W,Y		1°Anchor L,M,F,W,Y,I, Y,A
deleterious	A,G,P				G	G			

POSITION

	1	2	3	4	5	6	7	8	9 or C-terminus
B51 preferred	L,I,V,M,F,W,Y	¹⁰ Anchor P	F,W,Y	S,T,C	F,W,Y	G	G	F,W,Y	C-terminus ¹⁰ Anchor L,I,V,F,W, Y,A,M
deleterious	A,G,P,D,E,R,H,K, S,T,C		D,E	G	D,E,Q,N	G,D,E			
B5301 preferred	L,I,V,M,F,W,Y	¹⁰ Anchor P	F,W,Y	S,T,C	F,W,Y	L,I,V,M,F, W,Y	F,W,Y	¹⁰ Anchor I,M,F,W,Y, A,L,I,V	
deleterious	A,G,P,Q,N		G	R,H,K,Q,N	D,E				
B5401 preferred	F,W,Y	¹⁰ Anchor P	F,W,Y,L,I,V M	L,I,V,M	F,W,Y,A,P	A,L,I,V,M	F,W,Y,A,P	¹⁰ Anchor A,T,I,V,L, M,F,W,Y	
deleterious	G,P,Q,N,D,E		G,D,E,S,T,C	R,H,K,D,E	D,E	Q,N,D,G,E	D,E		

Italicized residues indicate less preferred or "tolerated" residues.
The information in Table II is specific for 9-mers unless otherwise specified.

Table III

MOTIFS	POSITION					
	1° anchor 1	2	3	4	5	6
DR4 preferred	F, M, Y, L, I, V, W	M	T		I	V, S, T, C, P, A, L, I, M
deleterious				W,		R, W, D, E
DR1 preferred	M, F, L, I, V, W, Y			P, A, M, Q		V, M, A, T, S, P, L, I, C
deleterious		C	C, H	F, D	C, W, D	G, D, E, D
DR7 preferred	M, F, L, I, V, W, Y	M	W	A		I, V, M, S, A, C, T, P, L
deleterious		C,		G,		G, R, D, N
DR Supermotif	M, F, L, I, V, W, Y					V, M, S, T, A, C, P, L, I
DR3 MOTIFS	1° anchor 1	2	3	4	5	1° anchor 6
motif a preferred	L, I, V, M, F, Y					
motif b preferred	L, I, V, M, F, A, Y					
						K, R, H

Italicized residues indicate less preferred or "tolerated" residues.

Table IV: HLA Class I Standard Peptide Binding Affinity.

ALLELE	STANDARD PEPTIDE	SEQUENCE (SEQ ID NO:)	STANDARD BINDING AFFINITY (nM)
A*0101	944.02	YLEPAIAKY	25
A*0201	941.01	FLPSDYFPSV	5.0
A*0202	941.01	FLPSDYFPSV	4.3
A*0203	941.01	FLPSDYFPSV	10
A*0205	941.01	FLPSDYFPSV	4.3
A*0206	941.01	FLPSDYFPSV	3.7
A*0207	941.01	FLPSDYFPSV	23
A*6802	1072.34	YVIKVSARV	8.0
A*0301	941.12	KVFPYALINK	11
A*1101	940.06	AVDLYHFLK	6.0
A*3101	941.12	KVFPYALINK	18
A*3301	1083.02	STLPETYVRR	29
A*6801	941.12	KVFPYALINK	8.0
A*2402	979.02	AYIDNYNKF	12
B*0702	1075.23	APRTLVL	5.5
B*3501	1021.05	PPFKYAAAF	7.2
B51	1021.05	PPFKYAAAF	5.5
B*5301	1021.05	PPFKYAAAF	9.3
B*5401	1021.05	PPFKYAAAF	10

Table V. HLA Class II Standard Peptide Binding Affinity.

Allele	Nomenclature	Standard Peptide	Sequence (SEQ ID NO:)	Binding Affinity (nM)
DRB1*0101	DR1	515.01	PKYVKQNTLKLAT	5.0
DRB1*0301	DR3	829.02	YKTIAFDEEARR	300
DRB1*0401	DR4w4	515.01	PKYVKQNTLKLAT	45
DRB1*0404	DR4w14	717.01	YARFQSQTTLKQKT	50
DRB1*0405	DR4w15	717.01	YARFQSQTTLKQKT	38
DRB1*0701	DR7	553.01	QYIKANSKFIGITE	25
DRB1*0802	DR8w2	553.01	QYIKANSKFIGITE	49
DRB1*0803	DR8w3	553.01	QYIKANSKFIGITE	1600
DRB1*0901	DR9	553.01	QYIKANSKFIGITE	75
DRB1*1101	DR5w11	553.01	QYIKANSKFIGITE	20
DRB1*1201	DR5w12	1200.05	EALIHQLKINPYVLS	298
DRB1*1302	DR6w19	650.22	QYIKANAKFIGITE	3.5
DRB1*1501	DR2w2β1	507.02	GRTQDENPVVHFFKNIV TPRTPPP	9.1
DRB3*0101	DR52a	511	NGQIGNDPNRDIL	470
DRB4*0101	DRw53	717.01	YARFQSQTTLKQKT	58
DRB5*0101	DR2w2β2	553.01	QYIKANSKFIGITE	20

Table VI

HLA-supertype	Verified*	Allele-specific HLA-supertype members	Predicted†
A1	A*0101, A*2501, A*2601, A*2602, A*3201	A*0102, A*2604, A*3601, A*4301, A*8001	
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, A*6901	A*0208, A*0210, A*0211, A*0212, A*0213	
A3	A*0301, A*1101, A*3101, A*3301, A*6801	A*0302, A*1102, A*2603, A*3302, A*3303, A*3401, A*3402, A*6601, A*6602, A*7401	
A24	A*2301, A*2402, A*3001	A*2403, A*2404, A*3002, A*3003	
B7	B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, B*7801	B*1511, B*4201, B*5901	
B27	B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, B*7301	B*2707, B*2708, B*3802, B*3903, B*3904, B*3905, B*4801, B*4802, B*1510, B*1518, B*1503	
B44	B*1801, B*1802, B*3701, B*4402, B*4403, B*4404, B*4001, B*4002, B*4006	B*4101, B*4301, B*4501, B*4901, B*5001	
B58	B*5701, B*5702, B*5801, B*5802, B*1516, B*1517		
B62	B*1501, B*1502, B*1513, B*5201	B*1301, B*1302, B*1504, B*1505, B*1506, B*1507, B*1515, B*1520, B*1521, B*1512, B*1514, B*1510	

- a. Verified alleles include alleles whose specificity has been determined by pool sequencing analysis, peptide binding assays, or by analysis of the sequences of CTL epitopes.
- b. Predicted alleles are alleles whose specificity is predicted on the basis of B and F pocket structure to overlap with the supertype specificity.

HICV A01 Super Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*0101
INQELVGV	1108	9	11	79	
RTTSYGVK	1295	10	11	79	
PMGFSYTRCF	2667	11	11	79	
PSVAATLGF	1261	9	14	100	
PLHGPTLLY	1821	11	11	79	
PQDDKLEF	1554	10	12	88	
PMGDSYTRCF	1554	10	12	88	
CVGSLQTF	1485	11	12	86	
RLKLSLF	2918	8	12	86	
RLAPLAY	1029	9	12	86	
RMWDMMHNR	317	10	12	86	
RLMLTTF	2375	0	12	86	
RLMLTTF	2375	9	12	86	
RVSCKALY	2621	9	14	100	
RLKLSLF	1554	8	12	86	
SVATLGF	1262	8	12	86	
SVATLGF	1262	11	14	100	
SVATLGF	1262	11	14	100	
TMANNEF	2590	9	11	79	
TLHGPTLLY	1822	10	11	79	
TLFNLGGW	1811	10	12	79	
TMANNEF	2309	10	12	79	
TMANNEF	1466	8	12	88	
TVSLQTF	1466	10	12	88	
VDTLCGF	132	9	12	86	
VLAALAY	1871	8	12	86	
VLDGNY	167	8	12	88	
VLDGNY	1552	9	11	79	
VLVOLAY	2639	8	11	79	
VMSSTGY	2639	10	14	100	
VMSSTGY	2639	10	14	100	
VMSSTGY	2639	9	11	79	
YKSDGDF	2848	11	11	79	
YKSDGDF	1106	11	11	79	
YKSDGDF	276	10	12	88	
YKGLSSVF		2			

0.0300

78

Table VIII
HCY A02 Super Motif with Binding Information

Conservancy	Freq.	Position	Sequence	A*0201	A*0202	A*0203	A*0206	A*0302
93	13	1994	AALRIHW					
86	12	1673	AALAAYCL					
79	11	1250	AAGGYVL					
79	11	1250	AAGGYVL					
79	11	1250	AAGGYVL					
79	11	147	AALAHIV					
79	11	147	AALAHIV					
100	14	1284	AALAHIV					
93	13	1284	AALAHIV					
86	12	1187	AALAHIV					
79	11	1187	AALAHIV					
79	11	1187	AALAHIV					
93	13	1690	AALAHIV					
86	12	1690	AALAHIV					
86	12	1690	AALAHIV					
100	14	150	AALAHIV					
100	14	150	AALAHIV					
96	12	1737	AALAHIV					
86	12	698	AALAHIV					
79	11	1698	AALAHIV					
79	11	1698	AALAHIV					
79	11	1698	AALAHIV					
86	12	1602	AALAHIV					
79	11	1251	AALAHIV					
79	11	1251	AALAHIV					
86	12	777	AALAHIV					
93	13	1265	AALAHIV					
79	11	1384	AALAHIV					
79	11	1384	AALAHIV					
100	14	1419	AALAHIV					
100	14	1419	AALAHIV					
79	11	1188	AALAHIV					
79	11	1188	AALAHIV					
79	11	1188	AALAHIV					
100	14	1917	AALAHIV					
100	14	1917	AALAHIV					
93	13	1903	AALAHIV					
79	11	1530	AALAHIV					
96	12	2941	AALAHIV					
86	12	739	AALAHIV					
79	11	1653	AALAHIV					

UICV A02 Super Motif with Binding Information

[illegible]

ICV AG2 Super Motif with Binding Information

Conservancy	Freq	Position	Sequence	A*0201	A*0202	A*0203	A*0208	A*6802
88	12	1670	GVLAALAA					
86	12	1670	GVLAALAYCL					
79	11	161	GVNATGNL	0.0001				
86	12	45	GVHATRTK					
100	14	2619	GVRCERMA					
100	14	2619	GVRCERMA					
100	14	2019	GVRCERMA	0.0002				
78	11	154	GVRCERMA	0.0001				
78	11	154	GVRCERMA					
100	14	1234	GVRCERMA					
100	14	1572	GVRCERMA					
86	12	696	GVRCERMA	0.0100	0.0014	0.5400	0.0027	0.0037
79	11	1719	GVRCERMA					
86	12	1769	GVRCERMA	0.3300	0.0004	0.1300	0.0280	0.0053
79	11	698	GVRCERMA					
79	11	222	GVRCERMA					
86	12	2855	GVRCERMA					
88	12	2855	GVRCERMA					
79	11	1910	GVRCERMA					
79	11	1910	GVRCERMA					
86	12	1923	GVRCERMA					
86	12	1923	GVRCERMA					
100	14	1925	GVRCERMA	0.0430	0.0300	2.0000	0.0049	0.0450
79	11	1925	GVRCERMA	0.0002				
79	11	1925	GVRCERMA					
79	11	1925	GVRCERMA					
86	12	1816	GVRCERMA	0.0430	0.0024	0.0190	0.0065	0.0038
86	12	1816	GVRCERMA					
86	12	1816	GVRCERMA					
86	12	1331	GVRCERMA					
86	12	1331	GVRCERMA					
93	13	1891	GVRCERMA					
93	13	1891	GVRCERMA					
93	13	1891	GVRCERMA					
93	13	1891	GVRCERMA					
79	11	2591	GVRCERMA	0.0210	0.0004	0.3700	0.0036	0.0130
100	14	1777	GVRCERMA	0.0068				
100	14	1777	GVRCERMA					
100	14	1777	GVRCERMA					
86	12	2250	GVRCERMA					
86	12	2250	GVRCERMA					
100	14	2816	GVRCERMA					
100	14	2816	GVRCERMA					
100	14	2816	GVRCERMA					
100	14	2816	GVRCERMA					
86	12	889	GVRCERMA					
86	12	889	GVRCERMA					

HCY A02 Super Model with Binding Information

Conservancy	Seq	Position	Sequence	A*0201	A*0202	A*0203	A*0206	A*0202
85	12	2240	LVVDEMGGN					
93	13	1629	LVVRLGAV					
79	11	133	LVGMPILV					
79	11	133	LVGMPILVGA					
88	12	2761	LVGTMVLY					
88	12	126	LVGPAAL					
86	12	126	LVGPAAL					
100	14	2180	LVGSPSHIT					
100	14	2180	LVGSPSHIT					
89	12	1852	LVGSKNOV					
83	13	1970	LVGDAHL					
83	13	1970	LVGDAHL					
79	11	2738	LVSMITPFSH					
79	11	2738	LVSMITPFSH					
88	12	1581	LVAYQATV					
88	12	1581	LVAYQATV					
79	11	1853	LVGLAGYGA					
88	12	1867	LVGGNAA					
88	12	1867	LVGGNAA					
88	12	1867	LVGGNAA					
88	12	1867	LVGGNAA					
100	14	1257	LVNPSVAA					
100	14	1257	LVNPSVAA					
100	14	1257	LVNPSVAA					
100	14	1257	LVNPSVAA					
100	14	1257	LVNPSVAA					
79	11	1884	LVNLPAIL					
79	11	1884	LVNLPAIL					
86	12	1137	LVTRHADV					
79	11	1137	LVTRHADV					
79	11	1137	LVTRHADV					
79	11	1897	LVGWCA					
79	11	1897	LVGWCA					
79	11	1897	LVGWCA					
79	11	1897	LVGWCA					
79	11	1897	LVGWCA					
79	11	1897	LVGWCA					
79	11	1897	LVGWCA					
86	12	1348	LVWDESK					
86	12	1348	LVWDESK					
100	14	2158	LVTPSHIT					
100	14	2158	LVTPSHIT					
100	14	2158	LVTPSHIT					
100	14	2158	LVTPSHIT					
83	13	322	LVMMNWSPT					

HCY A02 Super Motif with Binding Information

Conservancy	Freq.	Position	Sequence	A*0201	A*0202	A*0203	A*0206	A*6802
86	12	109	PTDPRRNL					
79	11	1473	PFTFRRL					
79	11	1473	PFTFTTT					
100	14	1236	PIGSGKST					
93	13	1236	PIGSGKTV					
86	12	1936	PIHVPESDA					
86	12	1936	PIHVPESDA					
79	11	1621	PLHGPTRL					
78	11	1621	PLHGPTRL					
78	11	2870	PTLWAKML					
78	11	2870	PTLWAKML					
78	11	2870	PTLWAKML					
78	11	2870	PTLWAKML					
100	14	1628	PTPLVRL					
93	13	1628	PTPLVRLGA					
93	13	1628	PTPLVRLGAV					
100	14	2657	PNSWLGNI	0.0001				
100	14	2657	PNSWLGNI	0.0001				
66	12	2657	PNSWLGNI					
79	11	2318	PVHGGTL	0.0004				
93	13	508	PVYCTPSV					
93	13	508	PVYCTPSV					
86	12	1340	QETAGARL					
86	12	1340	QETAGARL					
86	12	1340	QETAGARL					
86	12	1603	QAPPSKDM					
93	13	1555	QATVCNIA					
79	11	1555	QATVCNADA					
93	13	29	QNGGVYL	0.0015				
93	13	29	QNGGVYL					
86	12	336	QLLRQDA					
86	12	2184	QLPCEPQV	0.0002				
79	11	2210	QLSAPSLKA					
79	11	2210	QLSAPSLKA					
86	12	1465	QVCFSLTT					
86	12	1229	QVAFHAPT					
86	12	1186	RAVCTRGV					
79	11	1185	RAVCTRGVA	0.0001				
100	14	149	RAIAGWRV					
100	14	149	RAIAGWRV					
86	12	2733	RASGVLLT					
79	11	43	RLGWATRT					

LUCY A01 Super Motif with Binding Information

Conservancy	Freq.	Position	Sequence	A*0201	A*0202	A*0203	A*0206	A*5802
78	11	2916	RLQLSAFSL					
79	11	2611	RLVFLSL	0.0280	0.0055	0.0180	0.0002	0.0032
79	11	2611	RLNPLQGV					
79	11	1618	RLPTLUGRT	0.0890	0.0110	1.0000	0.0100	0.0050
86	12	1029	RLAPATA					
86	12	1347	RLVCLATA					
86	12	1347	RLWLATAT					
100	14	619	RLWHFCT					
86	12	317	RLWQDMM					
53	13	635	RLWQDMM					
86	12	2243	RLWQDMM					
86	12	2243	RLWQDMM					
86	12	2243	RLWQDMM					
86	12	2243	RLWQDMM					
78	11	1284	RLWQDMM					
78	11	1284	RLWQDMM					
100	14	2621	RLWQDMM					
86	12	2621	RLWQDMM					
86	12	2621	RLWQDMM					
86	12	2621	RLWQDMM					
86	12	2621	RLWQDMM					
79	11	2100	RLWQDMM	0.0001				
86	12	2252	RLWQDMM					
86	12	2252	RLWQDMM					
86	12	156	RLWQDMM					
86	12	156	RLWQDMM					
86	12	2033	RLWQDMM					
79	11	1655	RLWQDMM					
78	11	1655	RLWQDMM					
78	11	2212	RLWQDMM					
78	11	2212	RLWQDMM					
93	13	2207	RLWQDMM					
100	14	175	RLWQDMM					
86	12	175	RLWQDMM					
100	14	1470	RLWQDMM					
86	12	1470	RLWQDMM					
79	11	1470	RLWQDMM					
79	11	2926	RLWQDMM					
86	12	1051	RLWQDMM	0.0008				
100	14	2178	RLWQDMM	n 0003				
100	14	2178	RLWQDMM					
100	14	2178	RLWQDMM					
86	12	2163	RLWQDMM					
93	13	2209	RLWQDMM					
79	11	2209	RLWQDMM					
78	11	2209	RLWQDMM					

ICV A03 Super Motif (With Binding Information)

Conservancy	Freq.	Position	Sequence	A*9301	A*1101	A*3101	A*3301	A*6801
88	12	93	WGNGLSPR					
88	12	86	WLSPIGSR	0.0008	0.0005			
100	14	1920	WNRRLN/ASR	0.0530	0.0810			
79	11	877	WNRSLDTFK	0.0054	0.0005	0.0014	0.0420	0.0056
89	13	285	WNRSLDTR					
79	11	2830	YKSGEHL					
100	14	637	YKSGEHL					
86	12	1939	YYPESDAWF	0.0003	0.0001			
		112						

HUCV A24 Super Motif With Binding Information

[illegible]

HCV A24 Super Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A/2401
LGQTVL	1331	8	12	85	
IMAHNEIF	2391	8	12	85	
ITVSTNGIF	1296	9	12	86	
ITVSTNGIFL	1296	10	11	71	
NDQVLY	701	8	12	85	
NGQVLL	30	8	13	93	
NGVSTV	23	8	13	93	
NGVSTVDF	10	10	12	88	
LNILGQW	1813	8	12	88	
LEANLLW	2235	8	12	88	
LVNTGSL	414	8	11	79	
LVNTGSLW	170	8	12	86	
LVALLSC	1030	6	14	00	
LVNLLG	112	11	12	88	
LVNLLGSL	1897	11	13	88	
LVNLLGSL	36	9	13	85	
LVNLLGSL	87	11	11	79	
LVNLLGSL	2240	11	12	06	
LVNLLGSL	15	9	12	85	
LVNLLGSL	128	9	12	86	
LVNLLGSL	1570	11	12	86	
LVNLLGSL	1570	8	13	93	
LVNLLGSL	1570	9	13	93	
LVNLLGSL	2178	11	13	93	
LVNLLGSL	1570	8	11	79	
LVNLLGSL	1853	8	11	79	
LVNLLGSL	1687	9	12	79	
LVNLLGSL	1257	11	14	100	
LVNLLGSL	1884	8	11	79	
LVNLLGSL	1884	9	11	79	
LVNLLGSL	1897	10	11	79	
LVNLLGSL	1897	11	11	79	
LVNLLGSL	1897	11	11	79	
LVNLLGSL	2872	8	12	86	
LVNLLGSL	2872	11	12	86	
LVNLLGSL	2241	10	12	88	
LVNLLGSL	1135	11	11	79	
LVNLLGSL	2872	8	12	86	
LVNLLGSL	2179	8	14	100	
LVNLLGSL	1770	8	14	100	
LVNLLGSL	1770	11	14	100	
LVNLLGSL	838	10	13	100	
LVNLLGSL	1772	9	14	93	0.0270
LVNLLGSL	1772	9	14	100	0.0170

ICV A24 Super Motif With Binding Information

Sequence	Position	No. of Analog Res.	Sequence Frequency	Conservancy (%)	A ²⁴⁰¹
NLGSAWACL	1815	11	12	86	
NLRNTI	1282	9	79	86	
NLRNGL	1282	8	11	86	
NLRNGL	700	9	12	86	0.0001
NLRNGL	118	9	12	86	
NLRNGL	2239	8	12	86	
NLRNGL	168	9	13	93	
NLRNGL	188	10	13	93	
NLRNGL	188	11	13	93	
NLRNGL	1460	10	12	86	
NLRNGL	418	9	13	93	
NLRNGL	1108	11	11	79	
NLRNGL	1108	9	11	79	
NLRNGL	551	8	12	79	
NLRNGL	1295	10	11	79	
NLRNGL	1295	11	11	79	
NLRNGL	2403	11	13	93	
NLRNGL	143	9	11	79	
NLRNGL	2697	11	11	79	
NLRNGL	1621	11	11	79	
NLRNGL	1621	9	11	79	
NLRNGL	1821	10	11	79	
NLRNGL	1821	11	11	79	
NLRNGL	1821	11	11	79	
NLRNGL	2870	8	11	79	
NLRNGL	2870	9	11	79	
NLRNGL	1658	10	11	79	
NLRNGL	1658	9	14	100	
NLRNGL	1554	9	12	86	
NLRNGL	1554	10	12	86	
NLRNGL	2857	9	14	100	
NLRNGL	2857	10	14	100	
NLRNGL	2857	11	12	86	
NLRNGL	2318	8	11	79	
NLRNGL	1732	9	12	86	
NLRNGL	1732	10	12	86	
NLRNGL	28	9	13	93	
NLRNGL	28	11	12	86	
NLRNGL	1465	11	14	100	
NLRNGL	1919	9	14	100	0.0480
NLRNGL	1778	9	14	100	0.0180
NLRNGL	2647	10	11	79	
NLRNGL	2647	11	11	79	
NLRNGL	2818	10	12	86	
NLRNGL	2818	11	11	79	0.0001
NLRNGL	2811	8	11	79	

Sequence	Position	Peptide No.	No. of Amino Acids	Sequence Frequency	Centersancy [%]
RLAPITAY	1629		9	12	86
RMAYOMMMH	317		8	12	86
RMAYOMMMHH	317		10	12	86
RLMLTIIF	2875		8	12	86
RLMLTIIF	2875		9	12	86
RMAYOMMHL	625		11	13	93
RMAYOMMHL	2621		8	14	100
RMAYOMMHL	2621		9	14	100
RMAYOMMHL	156		9	14	100
RMAYOMMHL	173		9	14	100
SFSELALL	173		10	14	100
SFSELALL	175		8	14	100
SFSELALL	175		11	12	86
SFSELALL	175		12	12	86
SFSELALL	175		13	12	86
SFSELALL	175		14	12	86
SFSELALL	175		15	12	86
SFSELALL	175		16	12	86
SFSELALL	175		17	12	86
SFSELALL	175		18	12	86
SFSELALL	175		19	12	86
SFSELALL	175		20	12	86
SFSELALL	175		21	12	86
SFSELALL	175		22	12	86
SFSELALL	175		23	12	86
SFSELALL	175		24	12	86
SFSELALL	175		25	12	86
SFSELALL	175		26	12	86
SFSELALL	175		27	12	86
SFSELALL	175		28	12	86
SFSELALL	175		29	12	86
SFSELALL	175		30	12	86
SFSELALL	175		31	12	86
SFSELALL	175		32	12	86
SFSELALL	175		33	12	86
SFSELALL	175		34	12	86
SFSELALL	175		35	12	86
SFSELALL	175		36	12	86
SFSELALL	175		37	12	86
SFSELALL	175		38	12	86
SFSELALL	175		39	12	86
SFSELALL	175		40	12	86
SFSELALL	175		41	12	86
SFSELALL	175		42	12	86
SFSELALL	175		43	12	86
SFSELALL	175		44	12	86
SFSELALL	175		45	12	86
SFSELALL	175		46	12	86
SFSELALL	175		47	12	86
SFSELALL	175		48	12	86
SFSELALL	175		49	12	86
SFSELALL	175		50	12	86
SFSELALL	175		51	12	86
SFSELALL	175		52	12	86
SFSELALL	175		53	12	86
SFSELALL	175		54	12	86
SFSELALL	175		55	12	86
SFSELALL	175		56	12	86
SFSELALL	175		57	12	86
SFSELALL	175		58	12	86
SFSELALL	175		59	12	86
SFSELALL	175		60	12	86
SFSELALL	175		61	12	86
SFSELALL	175		62	12	86
SFSELALL	175		63	12	86
SFSELALL	175		64	12	86
SFSELALL	175		65	12	86
SFSELALL	175		66	12	86
SFSELALL	175		67	12	86
SFSELALL	175		68	12	86
SFSELALL	175		69	12	86
SFSELALL	175		70	12	86
SFSELALL	175		71	12	86
SFSELALL	175		72	12	86
SFSELALL	175		73	12	86
SFSELALL	175		74	12	86
SFSELALL	175		75	12	86
SFSELALL	175		76	12	86
SFSELALL	175		77	12	86
SFSELALL	175		78	12	86
SFSELALL	175		79	12	86
SFSELALL	175		80	12	86
SFSELALL	175		81	12	86
SFSELALL	175		82	12	86
SFSELALL	175		83	12	86
SFSELALL	175		84	12	86
SFSELALL	175		85	12	86

HCY A24 Super Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*2401
TYSTGRL	1297	8	13	93	
TYSTGRL	1297	9	12	86	0.0230
TYSTGRL	1366	8	13	93	
YDTCDE	1324	8	12	86	
YDTCDE	1324	9	12	86	
YDTCDE	1324	10	12	86	
YDTCDE	1324	11	12	86	
YDTCDE	1324	12	12	86	
YDTCDE	1324	13	12	86	
YDTCDE	1324	14	12	86	
YDTCDE	1324	15	12	86	
YDTCDE	1324	16	12	86	
YDTCDE	1324	17	12	86	
YDTCDE	1324	18	12	86	
YDTCDE	1324	19	12	86	
YDTCDE	1324	20	12	86	
YDTCDE	1324	21	12	86	
YDTCDE	1324	22	12	86	
YDTCDE	1324	23	12	86	
YDTCDE	1324	24	12	86	
YDTCDE	1324	25	12	86	
YDTCDE	1324	26	12	86	
YDTCDE	1324	27	12	86	
YDTCDE	1324	28	12	86	
YDTCDE	1324	29	12	86	
YDTCDE	1324	30	12	86	
YDTCDE	1324	31	12	86	
YDTCDE	1324	32	12	86	
YDTCDE	1324	33	12	86	
YDTCDE	1324	34	12	86	
YDTCDE	1324	35	12	86	
YDTCDE	1324	36	12	86	
YDTCDE	1324	37	12	86	
YDTCDE	1324	38	12	86	
YDTCDE	1324	39	12	86	
YDTCDE	1324	40	12	86	
YDTCDE	1324	41	12	86	
YDTCDE	1324	42	12	86	
YDTCDE	1324	43	12	86	
YDTCDE	1324	44	12	86	
YDTCDE	1324	45	12	86	
YDTCDE	1324	46	12	86	
YDTCDE	1324	47	12	86	
YDTCDE	1324	48	12	86	
YDTCDE	1324	49	12	86	
YDTCDE	1324	50	12	86	
YDTCDE	1324	51	12	86	
YDTCDE	1324	52	12	86	
YDTCDE	1324	53	12	86	
YDTCDE	1324	54	12	86	
YDTCDE	1324	55	12	86	
YDTCDE	1324	56	12	86	
YDTCDE	1324	57	12	86	
YDTCDE	1324	58	12	86	
YDTCDE	1324	59	12	86	
YDTCDE	1324	60	12	86	
YDTCDE	1324	61	12	86	
YDTCDE	1324	62	12	86	
YDTCDE	1324	63	12	86	
YDTCDE	1324	64	12	86	
YDTCDE	1324	65	12	86	
YDTCDE	1324	66	12	86	
YDTCDE	1324	67	12	86	
YDTCDE	1324	68	12	86	
YDTCDE	1324	69	12	86	
YDTCDE	1324	70	12	86	
YDTCDE	1324	71	12	86	
YDTCDE	1324	72	12	86	
YDTCDE	1324	73	12	86	
YDTCDE	1324	74	12	86	
YDTCDE	1324	75	12	86	
YDTCDE	1324	76	12	86	
YDTCDE	1324	77	12	86	
YDTCDE	1324	78	12	86	
YDTCDE	1324	79	12	86	
YDTCDE	1324	80	12	86	
YDTCDE	1324	81	12	86	
YDTCDE	1324	82	12	86	
YDTCDE	1324	83	12	86	
YDTCDE	1324	84	12	86	
YDTCDE	1324	85	12	86	
YDTCDE	1324	86	12	86	
YDTCDE	1324	87	12	86	
YDTCDE	1324	88	12	86	
YDTCDE	1324	89	12	86	
YDTCDE	1324	90	12	86	
YDTCDE	1324	91	12	86	
YDTCDE	1324	92	12	86	
YDTCDE	1324	93	12	86	
YDTCDE	1324	94	12	86	
YDTCDE	1324	95	12	86	
YDTCDE	1324	96	12	86	
YDTCDE	1324	97	12	86	
YDTCDE	1324	98	12	86	
YDTCDE	1324	99	12	86	
YDTCDE	1324	100	12	86	
YDTCDE	1324	101	12	86	
YDTCDE	1324	102	12	86	
YDTCDE	1324	103	12	86	
YDTCDE	1324	104	12	86	
YDTCDE	1324	105	12	86	
YDTCDE	1324	106	12	86	
YDTCDE	1324	107	12	86	
YDTCDE	1324	108	12	86	
YDTCDE	1324	109	12	86	
YDTCDE	1324	110	12	86	
YDTCDE	1324	111	12	86	
YDTCDE	1324	112	12	86	
YDTCDE	1324	113	12	86	
YDTCDE	1324	114	12	86	
YDTCDE	1324	115	12	86	
YDTCDE	1324	116	12	86	
YDTCDE	1324	117	12	86	
YDTCDE	1324	118	12	86	
YDTCDE	1324	119	12	86	
YDTCDE	1324	120	12	86	
YDTCDE	1324	121	12	86	
YDTCDE	1324	122	12	86	
YDTCDE	1324	123	12	86	
YDTCDE	1324	124	12	86	
YDTCDE	1324	125	12	86	
YDTCDE	1324	126	12	86	
YDTCDE	1324	127	12	86	
YDTCDE	1324	128	12	86	
YDTCDE	1324	129	12	86	
YDTCDE	1324	130	12	86	
YDTCDE	1324	131	12	86	
YDTCDE	1324	132	12	86	
YDTCDE	1324	133	12	86	
YDTCDE	1324	134	12	86	
YDTCDE	1324	135	12	86	
YDTCDE	1324	136	12	86	
YDTCDE	1324	137	12	86	
YDTCDE	1324	138	12	86	
YDTCDE	1324	139	12	86	
YDTCDE	1324	140	12	86	
YDTCDE	1324	141	12	86	
YDTCDE	1324	142	12	86	
YDTCDE	1324	143	12	86	
YDTCDE	1324	144	12	86	
YDTCDE	1324	145	12	86	
YDTCDE	1324	146	12	86	
YDTCDE	1324	147	12	86	
YDTCDE	1324	148	12	86	
YDTCDE	1324	149	12	86	
YDTCDE	1324	150	12	86	
YDTCDE	1324	151	12	86	
YDTCDE	1324	152	12	86	
YDTCDE	1324	153	12	86	
YDTCDE	1324	154	12	86	
YDTCDE	1324	155	12	86	
YDTCDE	1324	156	12	86	
YDTCDE	1324	157	12	86	
YDTCDE	1324	158	12	86	
YDTCDE	1324	159	12	86	
YDTCDE	1324	160	12	86	
YDTCDE	1324	161	12	86	
YDTCDE	1324	162	12	86	
YDTCDE	1324	163	12	86	
YDTCDE	1324	164	12	86	
YDTCDE	1324	165	12	86	
YDTCDE	1324	166	12	86	
YDTCDE	1324	167	12	86	
YDTCDE	1324	168	12	86	
YDTCDE	1324	169	12	86	
YDTCDE	1324	170	12	86	
YDTCDE	1324	171	12	86	
YDTCDE	1324	172	12	86	
YDTCDE	1324	173	12	86	
YDTCDE	1324	174	12	86	
YDTCDE	1324	175	12	86	
YDTCDE	1324	176	12	86	
YDTCDE	1324	177	12	86	
YDTCDE	1324	178	12	86	
YDTCDE	1324	179	12	86	
YDTCDE	1324	180	12	86	
YDTCDE	1324	181	12	86	
YDTCDE	1324	182	12	86	
YDTCDE	1324	183	12	86	
YDTCDE	1324	184	12	86	
YDTCDE	1324	185	12	86	
YDTCDE	1324	186	12	86	
YDTCDE	1324	187	12	86	
YDTCDE	1324	188	12	86	
YDTCDE	1324	189	12	86	
YDTCDE	1324	190	12	86	
YDTCDE	1324	191	12	86	
YDTCDE	1324	192	12	86	
YDTCDE	1324	193	12	86	
YDTCDE	1324	194	12	86	
YDTCDE	1324	195	12	86	
YDTCDE	1324	196	12	86	
YDTCDE	1324	197	12	86	
YDTCDE	1324	198	12	86	
YDTCDE	1324	199	12	86	
YDTCDE	1324	200	12	86	
YDTCDE	1324	201	12	86	
YDTCDE	1324	202	12	86	
YDTCDE	1324	203	12	86	
YDTCDE	1324	204	12	86	
YDTCDE	1324	205	12	86	
YDTCDE	1324	206	12	86	
YDTCDE	1324	207	12	86	
YDTCDE	1324	208	12	86	
YDTCDE	1324	209	12	86	
YDTCDE	1324	210	12	86	
YDTCDE	1324	211	12	86	
YDTCDE	1324	212	12	86	
YDTCDE	1324	213	12	86	
YDTCDE	1324	214	12	86	
YDTCDE	1324	215	12	86	
YDTCDE	1324	216	12	86	
YDTCDE	1324	217	12	86	
YDTCDE	1324	218	12	86	
YDTCDE	1324	219	12	86	
YDTCDE	1324	220	12	86	
YDTCDE	1324	221	12	86	
YDTCDE	1324	222	12	86	
YDTCDE	1324	223	12	86	
YDTCDE	1324	224	12	86	
YDTCDE	1324	225	12	86	
YDTCDE	1324	226	12	86	
YDTCDE	1324	227	12	86	
YDTCDE	1324	228	12	86	
YDTCDE	1324	229	12	86	
YDTCDE	1324	230	12	86	
YDTCDE	1324	231	12	86	
YDTCDE	1324	232	12	86	
YDTCDE	1324	233	12	86	
YDTCDE	1324	234	12	86	
YDTCDE	1324	235	12	86	
YDTCDE	1324	236	12	86	
YDTCDE	1324	237	12	86	
YDTCDE	1324	238	12	86	
YDTCDE	1324	239	12	86	
YDTCDE	1324	240	12	86	
YDTCDE	1324	241	12	86	
YDTCDE	1324	242	12	86	
YDTCDE	1324	243	12	86	
YDTCDE	1324	244	12	86	
YDTCDE	1324	245	12	86	
YDTCDE	1324	246	12	86	
YDTCDE	1324	247	12	86	
YDTCDE	1324	248	12	86	
YDTCDE	1324	249	12	86	
YDTCDE	1324	250	12	86	

ICV 207 Super Motif (with Binding Information)

Conservancy	Freq.	Position	Sequence	B*0702	B*3501	B*5101	B*5301	B*5401
86	12	78	CGGKRWLY	0.0001	0.0011	0.0002	0.0001	0.0002
87	13	239	YDQVY	0.2300	0.0002	0.0001	0.0001	0.0002
79	11	239	PHQVYK	0.0001	0.0002	0.0001	0.0001	0.0002
93	13	1893	SPGALVGV	0.0001	0.0002	0.0002	0.1200	0.0002
79	11	1893	SPGALVGV	0.0130	0.0001	0.0018	0.0001	0.0003
79	11	2831	SPQENRV	0.0007	0.0001	0.0001	0.0002	0.0007
79	11	2831	SPQENRVA	0.0003	0.0001	0.0001	0.0002	0.0007
78	11	2849	SPQENR	0.0027	0.0002	0.0002	0.0001	0.0002
78	11	2849	SPQENR	0.1200	0.0002	0.0002	0.0001	0.0002
79	11	189	SPKSPSPV	0.0001	0.0002	0.0005	0.0001	0.0002
79	11	189	SPKSPSPV	0.0001	0.0001	0.0001	0.0001	0.0001
86	12	1835	SPTHVFESDA	0.0001	0.0001	0.0001	0.0001	0.0003
86	12	1875	TPCGSGL	0.0028	0.0001	0.0002	0.0001	0.0003
79	11	1128	TPCTGSSSL	0.0005	0.0005	0.0002	0.0001	0.0003
79	11	1128	TPCTGSSSL	0.0001	0.0001	0.0001	0.0001	0.0001
86	12	223	TPQVPCV	0.0001	0.0001	0.0001	0.0001	0.0001
93	13	1550	TPGLPQCDL	0.0001	0.0001	0.0001	0.0002	0.2300
93	13	1827	TPLLRGA	0.0083	0.0001	0.0001	0.0002	0.0001
86	12	2850	TPNSALGN	0.0110	0.0001	0.0009	0.0001	0.0009
86	12	2858	TPNSALGHI	0.0001	0.0001	0.0003	0.0005	0.0009
86	12	1840	VPESDAAA	0.0022	0.0001	0.0001	0.0001	0.0003
86	12	1840	VPESDAARV	0.0001	0.0001	0.0010	0.0001	0.0003
86	12	799	WPLLLLL	0.0001	0.0001	0.0001	0.0001	0.0001
100	14	618	YPTLWLY	0.0021	0.0001	0.0001	0.0001	0.0001

76

Table XIII

[illegible]

HCV B27 Super Motif

[illegible]

HCV B27 Super Motif

[illegible]

UCV 1027 Super-Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
YKSLVNFSA 136	1254	11	14	100

HCV B58 Super Motif Table XIII

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
ANLRRPV	1604	8	13	93
ALAAAYCL	1673	8	13	86
AAQGYVL	1250	8	11	79
ANLIGGA	1264	8	14	100
AAVCTRGV	1187	8	12	88
ASLMFTA	1793	8	11	79
ASSASQL	2204	8	14	100
ATLGGAY	1265	8	14	100
CSLGL	1778	8	10	100
CSGANGDL	1316	8	12	86
CSNSNSVA	2619	8	14	100
CTCGSSSL	1128	8	11	79
CTRRVAKA	1180	8	11	79
DTAACDGI	884	8	12	86
DTLTCGFA	124	8	12	86
EALENLY	124	8	12	86
EAENLY	750	8	11	100
EAENLY	274	8	11	100
ESDAARV	1842	8	12	86
ETAGARV	1342	8	12	86
ETTMRSFV	1207	8	12	86
FACLAGYI	130	8	13	93
FAGRNRY	1927	8	14	100
FSPILLAL	174	8	14	100
FSPILLAL	2782	8	11	100
FTETMTRY	2782	8	14	100
FTFSPVAV	512	8	13	83
GAGVAGAL	1861	8	12	86
GAHNGILA	350	8	12	88
GALVGVV	1985	8	11	79
GARLVLA	1345	8	12	88
GGGGITRV	1238	8	13	88
GGGGLTV	1181	8	13	88
GGGGLTV	1188	8	12	88
GGSGGCVL	2841	8	11	79
GGSGGCVY	2841	8	11	79
GTFPNAY	2083	8	11	79
HSTSPGEI	2928	8	11	79
HIPNNSWL	2955	8	12	86
ISGIOTLA	1774	8	14	100
ITSCSSRW	2816	8	14	100
ITSCSSRW	1928	8	14	100
KSTVAPAA	1241	8	12	88
KSTVAPAA	1241	8	12	88
LAGTGARV	1857	8	11	79
LAHSRVVL	151	8	14	100
LAVAVPV	972	8	11	79
LSNP/SLKA	2211	8	11	79

HCY B58 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Consensus (%)
LSFGALVV	1892	8	13	93
LSTGHL	890	8	12	86
LTCGFQIL	126	8	12	86
LTHDAHF	1570	8	13	93
MSADLEVV	1654	8	11	79
NSMGLNI	2859	8	14	100
NTC/TQTV	1460	8	12	88
NKNGSWFI	416	8	13	93
PALLSPGA	1889	8	12	88
PALSTLGL	1890	8	13	93
PILWALNI	2870	8	11	79
PTLWALNI	1828	8	14	100
FTPLVRL	1595	8	13	93
QATVCARA	3019	8	14	100
RAAPRWFM	664	8	11	79
RSLSPLL	664	8	12	86
RSPLGV	115	8	11	79
SSKQVLA	2262	8	11	79
SSKQSLA	2268	8	12	86
STKVAAY	1242	8	12	86
STLPQNP	1764	8	14	100
STLPQNM	2653	8	12	86
STYGCLA	1299	8	12	86
TYAGDIL	995	8	12	86
TYAGDIL	1302	8	12	86
TYAGSPV	1208	8	12	86
TTMSGPF	2738	8	12	75
TTSGDITL	1864	8	11	79
VAGALVF	1136	8	13	66
VTRHADNI	1681	8	12	86
VSTWLV	1681	8	12	86
WAGHWNF	1766	8	12	86
WAKULM	388	8	14	100
WAKSPW	1768	8	12	86
WAKSPW	1768	8	11	79
WAKSPW	2895	8	13	93
YSERDL	1298	8	12	86
YSTGKRL	1298	8	11	79
YTMQDL	1106	9	16	114
AKKLDCTM	2758	9	11	79
AKOSTAVLV	1250	9	11	79
AKRLAHV	147	9	14	100
AKTLRGNV	1284	9	11	79
AKVLSKSL	1262	9	12	86
AKVLSKSL	2208	9	22	186
ATLGFQVM	1265	9	11	79
ATVCARQCA	1586	9	11	79
CAALLRHW	1903	9	13	93

HCV B58 Super Matif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
GWVETPA	1630	9	11	79
GFSELLA	172	9	14	100
GGGPRFA	1110	9	11	79
CTGGSSA	1120	9	12	86
CTGGSSA	1130	9	11	79
CTGGSSA	1140	9	11	79
CTGWSSTGF	555	2	11	79
DAGSNWYEL	1527	9	11	79
DYACEDII	994	8	12	86
DYICRSTV	2673	9	13	93
ETAGRLW	1342	9	12	86
ETNARSMF	1207	9	12	86
FIRLLALL	174	9	14	100
FIRLLALL	175	9	14	100
FTGLTHDA	1587	9	12	86
GAIVGALV	1801	9	12	86
GALVAFRM	1866	9	12	86
GALVAFKVA	1868	9	14	100
GANVMARL	1818	9	14	100
HSKCKCEL	1400	8	14	100
HTPGRCPOV	222	9	11	79
ITNIGADIAT	989	9	12	86
ITNIGADIAT	990	9	12	86
ITNIGADIAT	991	9	12	86
KALVADLV	1241	9	12	86
KSTNMPAY	1672	9	12	86
LALAALYC	1729	9	12	86
LAECQKQA	1729	9	12	86
LGLAYTSM	356	9	14	100
LAGTQAGVA	1657	9	11	79
LSAFSLUST	2922	9	11	79
LETPRQPA	1789	9	14	100
LTCQADLLA	126	9	24	171
LTPRQKQA	206	9	14	100
LTPRQKQA	1052	9	12	86
LTPRQKQA	1052	9	13	93
LTHIDANVL	1570	9	11	79
LITSGSMTL	2738	9	11	79
MANNEVFO	2592	9	12	86
MANVMDMMHW	318	9	12	86
NAVAYYEL	1418	9	13	93
NSLIRPRRM	2481	9	14	100
NSWGNHM	2859	9	24	171
NINPRQDQ	14	9	12	86
NINPRQDQ	116	9	13	93
NSVATLGE	126	9	14	100
PSVATLGE	1621	9	11	79
PTLHGPTPL	2870	9	11	79
PTLHWMAL	2870	9	11	79

HCV B58 Super Motif

Sequence	Position	No. of Analog Acids	Sequence Frequency	Conservancy [%]
GAETAGRL	1340	9	12	86
RAVCTHGV	1196	9	12	86
RAVCTHGV	1480	9	14	86
RAVCTHGV	1801	9	11	78
RAVAMREH	811	9	15	114
RSLSKLL	664	2	11	79
RSINLGM	115	9	12	86
SSASQSLA	2205	9	14	100
STKVFAYA	1242	9	12	86
STLPONPJ	1784	9	11	79
STWLVGV	1663	9	12	86
TAGARLVK	1343	9	12	86
TSSSNV	1250	9	14	100
TTMANNEV	2589	9	11	79
VAILGFGA	1263	9	14	100
VAGSHYQM	933	9	14	100
VAYQATYA	1592	9	12	86
VAYYRGDY	1490	9	14	100
VSTLQANM	1252	9	12	86
VTDQDEL	1483	9	12	86
WRRHNNR	1768	9	12	86
YAGGVKVL	1249	8	11	79
YAPTLWARM	2668	9	14	100
YSPGSHV	2930	9	11	78
YSPGSHV	2648	8	11	78
YTHKAKA	1186	9	12	86
YVYDQVL	1195	9	12	79
YVYDQVL	1250	10	11	79
AAQGVKVL	1264	10	28	186
AAILGFAYM	1264	10	28	186
ASLRVFEM	2787	10	12	86
ASSASQSLA	2204	10	14	100
ATGNPCSF	165	10	13	93
CFPSRLAL	1172	10	14	100
CHGSSSLYL	1128	10	11	86
DMKQCNH	1250	10	11	120
DMKQCNH	1434	10	12	86
DILTOSFAL	1254	10	12	86
ENILMRCEM	2237	10	24	171
ETAGARLVK	1342	10	12	86
FAGLKGTHL	1309	10	11	79
FTEAMTRISA	2792	10	14	100
GAGVAGVAGV	1483	10	14	100
GATVACDQDI	992	10	12	86
GAGVAGVAV	1661	10	12	86
GALVGVNCA	1685	10	11	79

HCY B58 Super Methyl

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
GARLVVLATA	13145	10	11	79
GRKQKQKQK	13116	10	14	100
RSQKSTVPA	13116	10	12	88
GTALQDAETTA	13335	10	14	100
HSKKQKDELA	14000	10	14	100
IAFASRGQHH	13925	10	14	100
ISGIOTLAGL	17774	10	12	88
ITRSESNNV	22560	10	14	100
ITSCSNNSV	22818	10	14	100
ITSTSTGKPL	13245	10	14	100
ITSTSTGKPL	13245	10	12	88
LAQKQKQKQK	13305	10	12	88
LAQKQKQKQK	13305	10	11	79
LALPFRATAM	17289	10	12	88
LALPFRATAM	606	10	12	88
LEPGALVGV	1832	10	13	93
LEPGSRPSW	88	10	11	79
LSRAPRRPM	3017	10	14	100
LSRAPRRPM	3017	10	11	79
LTHRTKYM	1842	10	16	114
NTQVTDVDF	1460	10	12	88
PAISPGALV	1888	10	12	88
PAISGLJHL	688	10	12	88
PARLVFPL	2609	10	11	79
PSMDMMKGL	1607	10	13	93
PTGGSRSTV	1836	10	13	93
PTGGSRSTV	1836	10	12	88
PTLWGRVLA	1821	10	11	79
PTLWARMIL	2370	10	22	157
PTLLYRLGA	1628	10	13	93
QAEATGARLV	1340	10	12	88
QAEATGARLV	1603	10	24	171
QATYCNAGQA	1595	10	11	79
QATYCNAGQA	27787	10	16	114
RANGLDCTM	1465	10	13	93
RANGLDCTM	1465	10	11	79
RALHGHVRA	148	10	14	100
SASQLSAPSL	2207	10	13	93
STKVPATYAA	1242	10	11	79
STWLVGSL	1663	10	12	88
TAGARLVULA	1343	10	12	88
TARHTPNSW	2852	10	11	79
TGCSNNSWVA	2317	10	13	93
TNLTITPNSL	17774	10	14	100
TNLTITPNSL	1662	10	12	88
TNLTITPNSL	1662	10	13	93
TINAKNEVF	2688	10	11	79
TILPALSTGL	685	10	11	79

ICV B58 Super Match

Sequence	Position	No. of Ancho Addit	Sequence Frequency	Consistency (%)
WANTGFRAY	1263	10	14	100
WNTGFRAY	1507	10	16	114
WTRADAPV	1158	10	11	79
WAFSYWPL	76	10	12	86
WARMILATH	2873	10	12	86
WAPDYNPL	2297	10	11	79
WAOBYWLV	1243	10	11	79
WSPENRVA	2530	10	11	79
YSPGJNRL	2648	10	11	79
ANALNIGVR	140	11	11	79
ASLNTFRM	218	11	12	86
ASLNTFRM	1187	11	11	79
ASHPIECOM	1717	11	14	100
ASQISAPSKA	2208	11	11	79
CAROMPPSW	1559	11	11	79
CSFSFLAL	172	11	14	100
CTGSSOLV	1128	11	11	79
CTGSSOLV	1135	11	11	79
DARCACLVMA	1351	11	11	79
DTLCSFQML	723	11	16	114
ETAGARLVLA	1342	11	24	171
FALMGIPLV	130	11	12	86
FLHSHSPGE	2925	11	11	79
FTLTHIDHF	1507	11	11	79
FTLPALETOL	1507	11	12	93
GAUWAGVLA	892	11	11	79
GAUWAGVLF	1861	11	12	86
QALVGVVICA	1895	11	12	86
GAUWGNRLA	1818	11	11	79
GSCKSTVPA	1238	11	14	100
HSNNKCEJAA	1400	11	12	86
HSYSPGEMV	2928	11	14	100
HFWNWDON	2855	11	11	79
ITNESGKRV	2516	11	12	86
ITNESGKRV	2516	11	12	86
ITSTYKVELA	1296	11	14	100
KSITVPAYAA	1241	11	11	79
LAGGSCSGAY	1305	11	11	79
LAGYGAHADA	1857	11	11	79
LSNSLRHNM	2479	11	14	100
LSFGALVGV	1982	11	11	79
LSFGALVGV	1982	11	12	86
LSFGALVGV	1982	11	13	93
LSFGALVGV	1982	11	13	93
NAVATYGLDV	1418	11	13	93
NTNRPQDQF	14	11	11	79

ICV B58 Super-Mohr

Sequence	Position	No. of Amino Acids	Sequence Frequency	Consistency (%)
PALLSGALVY	1689	11	12	86
PTVATLGGCA	1261	11	14	100
PTVPTTSSRL	109	11	12	86
PTVPTTSSDA	1836	11	12	86
PTLHGPTLLY	1621	11	11	78
PTLLYLRGAV	1628	11	13	86
QAEETAGARLV	1340	11	12	78
QAPFSSQDAW	1603	11	12	86
QVYDFSLPTF	1465	11	12	86
RSZPTPTPEPI	165	11	13	93
RSZPTPTPEV	1655	11	11	78
SSQSLVLRV	2208	11	13	93
SSQSLVLRSL	1132	11	13	93
STWNLGGWLA	1663	11	12	86
TARHTPMSWL	2862	11	12	86
TSLTGRDNKV	1050	11	12	78
TSVWNLGGVL	1825	11	12	86
TSVWNLGGV	895	11	12	86
TSVWNLGGV	1343	11	11	79
VAKTLRGANVA	1864	11	11	79
VASALVAFKVA	1944	11	28	188
VAVEPTWFSQ	974	11	14	100
VAYDQVTCARA	1592	11	12	86
VAYYRELDSV	1420	11	12	78
VYTWLWLGCV	1581	11	14	100
WAKSTVAPKLV	72	11	12	86
WAKSTVAPKLV	2873	11	12	86
YAGQTPKATF	1249	11	11	78
YATGAI PQCSF	154	11	12	86
YTMQDLGVW	1106	11	11	78

288

HCV B62 Super Molif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
ALLSPGAL	1890	8	13	83
ALNKEVIV	1891	8	13	100
ALNKEVIV	1892	8	13	88
ALGLDIA	1893	8	12	88
ALGLDIA	1894	8	11	78
ALGLDIA	1895	8	11	88
ALGLDIA	1896	8	12	88
ALGLDIA	1897	8	11	78
ALGLDIA	1898	8	11	100
ALGLDIA	1899	8	14	100
ALGLDIA	1900	8	11	79
ALGLDIA	1901	8	11	79
ALGLDIA	1902	8	11	79
ALGLDIA	1903	8	11	79
ALGLDIA	1904	8	11	79
ALGLDIA	1905	8	11	79
ALGLDIA	1906	8	11	79
ALGLDIA	1907	8	11	79
ALGLDIA	1908	8	11	79
ALGLDIA	1909	8	11	79
ALGLDIA	1910	8	11	79
ALGLDIA	1911	8	11	79
ALGLDIA	1912	8	12	86
ALGLDIA	1913	8	12	86
ALGLDIA	1914	8	12	86
ALGLDIA	1915	8	12	86
ALGLDIA	1916	8	12	86
ALGLDIA	1917	8	12	86
ALGLDIA	1918	8	12	86
ALGLDIA	1919	8	12	86
ALGLDIA	1920	8	12	86
ALGLDIA	1921	8	12	86
ALGLDIA	1922	8	12	86
ALGLDIA	1923	8	12	86
ALGLDIA	1924	8	12	86
ALGLDIA	1925	8	12	86
ALGLDIA	1926	8	12	86
ALGLDIA	1927	8	12	86
ALGLDIA	1928	8	12	86
ALGLDIA	1929	8	12	86
ALGLDIA	1930	8	12	86
ALGLDIA	1931	8	12	86
ALGLDIA	1932	8	12	86
ALGLDIA	1933	8	12	86
ALGLDIA	1934	8	12	86
ALGLDIA	1935	8	12	86
ALGLDIA	1936	8	12	86
ALGLDIA	1937	8	12	86
ALGLDIA	1938	8	12	86
ALGLDIA	1939	8	12	86
ALGLDIA	1940	8	12	86
ALGLDIA	1941	8	12	86
ALGLDIA	1942	8	12	86
ALGLDIA	1943	8	12	86
ALGLDIA	1944	8	12	86
ALGLDIA	1945	8	12	86
ALGLDIA	1946	8	12	86
ALGLDIA	1947	8	12	86
ALGLDIA	1948	8	12	86
ALGLDIA	1949	8	12	86
ALGLDIA	1950	8	12	86
ALGLDIA	1951	8	12	86
ALGLDIA	1952	8	12	86
ALGLDIA	1953	8	12	86
ALGLDIA	1954	8	12	86
ALGLDIA	1955	8	12	86
ALGLDIA	1956	8	12	86
ALGLDIA	1957	8	12	86
ALGLDIA	1958	8	12	86
ALGLDIA	1959	8	12	86
ALGLDIA	1960	8	12	86
ALGLDIA	1961	8	12	86
ALGLDIA	1962	8	12	86
ALGLDIA	1963	8	12	86
ALGLDIA	1964	8	12	86
ALGLDIA	1965	8	12	86
ALGLDIA	1966	8	12	86
ALGLDIA	1967	8	12	86
ALGLDIA	1968	8	12	86
ALGLDIA	1969	8	12	86
ALGLDIA	1970	8	12	86
ALGLDIA	1971	8	12	86
ALGLDIA	1972	8	12	86
ALGLDIA	1973	8	12	86
ALGLDIA	1974	8	12	86
ALGLDIA	1975	8	12	86
ALGLDIA	1976	8	12	86
ALGLDIA	1977	8	12	86
ALGLDIA	1978	8	12	86
ALGLDIA	1979	8	12	86
ALGLDIA	1980	8	12	86
ALGLDIA	1981	8	12	86
ALGLDIA	1982	8	12	86
ALGLDIA	1983	8	12	86
ALGLDIA	1984	8	12	86
ALGLDIA	1985	8	12	86
ALGLDIA	1986	8	12	86
ALGLDIA	1987	8	12	86
ALGLDIA	1988	8	12	86
ALGLDIA	1989	8	12	86
ALGLDIA	1990	8	12	86
ALGLDIA	1991	8	12	86
ALGLDIA	1992	8	12	86
ALGLDIA	1993	8	12	86
ALGLDIA	1994	8	12	86
ALGLDIA	1995	8	12	86
ALGLDIA	1996	8	12	86
ALGLDIA	1997	8	12	86
ALGLDIA	1998	8	12	86
ALGLDIA	1999	8	12	86
ALGLDIA	2000	8	12	86
ALGLDIA	2001	8	12	86
ALGLDIA	2002	8	12	86
ALGLDIA	2003	8	12	86
ALGLDIA	2004	8	12	86
ALGLDIA	2005	8	12	86
ALGLDIA	2006	8	12	86
ALGLDIA	2007	8	12	86
ALGLDIA	2008	8	12	86
ALGLDIA	2009	8	12	86
ALGLDIA	2010	8	12	86
ALGLDIA	2011	8	12	86
ALGLDIA	2012	8	12	86
ALGLDIA	2013	8	12	86
ALGLDIA	2014	8	12	86
ALGLDIA	2015	8	12	86
ALGLDIA	2016	8	12	86
ALGLDIA	2017	8	12	86
ALGLDIA	2018	8	12	86
ALGLDIA	2019	8	12	86
ALGLDIA	2020	8	12	86
ALGLDIA	2021	8	12	86
ALGLDIA	2022	8	12	86
ALGLDIA	2023	8	12	86
ALGLDIA	2024	8	12	86
ALGLDIA	2025	8	12	86
ALGLDIA	2026	8	12	86
ALGLDIA	2027	8	12	86
ALGLDIA	2028	8	12	86
ALGLDIA	2029	8	12	86
ALGLDIA	2030	8	12	86
ALGLDIA	2031	8	12	86
ALGLDIA	2032	8	12	86
ALGLDIA	2033	8	12	86
ALGLDIA	2034	8	12	86
ALGLDIA	2035	8	12	86
ALGLDIA	2036	8	12	86
ALGLDIA	2037	8	12	86
ALGLDIA	2038	8	12	86
ALGLDIA	2039	8	12	86
ALGLDIA	2040	8	12	86
ALGLDIA	2041	8	12	86
ALGLDIA	2042	8	12	86
ALGLDIA	2043	8	12	86
ALGLDIA	2044	8	12	86
ALGLDIA	2045	8	12	86
ALGLDIA	2046	8	12	86
ALGLDIA	2047	8	12	86
ALGLDIA	2048	8	12	86
ALGLDIA	2049	8	12	86
ALGLDIA	2050	8	12	86
ALGLDIA	2051	8	12	86
ALGLDIA	2052	8	12	86
ALGLDIA	2053	8	12	86
ALGLDIA	2054	8	12	86
ALGLDIA	2055	8	12	86
ALGLDIA	2056	8	12	86
ALGLDIA	2057	8	12	86
ALGLDIA	2058	8	12	86
ALGLDIA	2059	8	12	86
ALGLDIA	2060	8	12	86
ALGLDIA	2061	8	12	86
ALGLDIA	2062	8	12	86
ALGLDIA	2063	8	12	86
ALGLDIA	2064	8	12	86
ALGLDIA	2065	8	12	86
ALGLDIA	2066	8	12	86
ALGLDIA	2067	8	12	86
ALGLDIA	2068	8	12	86
ALGLDIA	2069	8	12	86
ALGLDIA	2070	8	12	86
ALGLDIA	2071	8	12	86
ALGLDIA	2072	8	12	86
ALGLDIA	2073	8	12	86
ALGLDIA	2074	8	12	86
ALGLDIA	2075	8	12	86
ALGLDIA	2076	8	12	86
ALGLDIA	2077	8	12	86
ALGLDIA	2078	8	12	86
ALGLDIA	2079	8	12	86
ALGLDIA	2080	8	12	86
ALGLDIA	2081	8	12	86
ALGLDIA	2082	8	12	86
ALGLDIA	2083	8	12	86
ALGLDIA	2084	8	12	86
ALGLDIA	2085	8	12	86
ALGLDIA	2086	8	12	86
ALGLDIA	2087	8	12	86
ALGLDIA	2088	8	12	86
ALGLDIA	2089	8	12	86
ALGLDIA	2090	8	12	86
ALGLDIA	2091	8	12	86
ALGLDIA	2092	8	12	86
ALGLDIA	2093	8	12	86
ALGLDIA	2094	8	12	86
ALGLDIA	2095	8	12	86
ALGLDIA	2096	8	12	86
ALGLDIA	2097	8	12	86
ALGLDIA	2098	8	12	86
ALGLDIA	2099	8	12	86
ALGLDIA	2100	8	12	86
ALGLDIA	2101	8	12	86
ALGLDIA	2102	8	12	86
ALGLDIA	2103	8	12	86
ALGLDIA	2104	8	12	86
ALGLDIA	2105	8	12	86
ALGLDIA	2106	8	12	86
ALGLDIA	2107	8	12	86
ALGLDIA	2108	8	12	86
ALGLDIA	2109	8	12	86
ALGLDIA	2110	8	12	86
ALGLDIA	2111	8	12	86
ALGLDIA	2112	8	12	86
ALGLDIA	2113	8	12	86
ALGLDIA	2114	8	12	86
ALGLDIA	2115	8	12	86
ALGLDIA	2116	8	12	86
ALGLDIA	2117	8	12	86
ALGLDIA	2118	8	12	86
ALGLDIA	2119	8	12	86
ALGLDIA	2120	8	12	86
ALGLDIA	2121	8	12	86
ALGLDIA	2122	8	12	86
ALGLDIA	2123	8	12	86
ALGLDIA	2124	8	12	86
ALGLDIA	2125	8	12	86
ALGLDIA	2126	8	12	86
ALGLDIA	2127	8	12	86
ALGLDIA	2128	8	12	86
ALGLDIA	2129	8	12	86
ALGLDIA	2130	8	12	86
ALGLDIA	2131	8	12	86
ALGLDIA	2132	8	12	86
ALGLDIA	2133	8	12	86
ALGLDIA	2134	8	12	86
ALGLDIA	2135	8	12	86
ALGLDIA	2136	8	12	86
ALGLDIA	2137	8	12	86
ALGLDIA	2138	8	12	86
ALGLDIA	2139	8	12	86
ALGLDIA	2140	8	12	86
ALGLDIA	2141	8	12	86
ALGLDIA	2142	8	12	86
ALGLDIA	2143	8	12	86
ALGLDIA	2144	8	12	86
ALGLDIA	2145	8	12	86
ALGLDIA	2146	8	12	86
ALGLDIA	2147	8	12	86
ALGLDIA	2148	8	12	86
ALGLDIA	2149	8	12	86
ALGLDIA	2150	8	12	86
ALGLDIA	2151	8	12	86
ALGLDIA	2152	8	12	86
ALGLDIA	2153	8	12	86
ALGLDIA	2154	8	12	86
ALGLDIA	2155	8	12	86
ALGLDIA	2156	8	12	86
ALGLDIA	2157	8	12	86
ALGLDIA	2158	8	12	86
ALGLDIA	2159	8	12	86
ALGLDIA	2160	8	12	86
ALGLDIA	2161	8	12	86
ALGLDIA	2162	8	12	86
ALGLDIA	2163	8	12	86
ALGLDIA	2164	8	12	86
ALGLDIA	2165	8	12	86
ALGLDIA	2166	8	12	86
ALGLDIA	2167	8	12	86
ALGLDIA	2168	8	12	86
ALGLDIA	2169	8	12	86
ALGLDIA	2170	8	12	86
ALGLDIA	2171	8	12	86
ALGLDIA	2172	8	12	86
ALGLDIA	2173	8	12	86
ALGLDIA	2174	8	12	86
ALGLDIA	2175	8	12	86
ALGLDIA	2176	8	12	86
ALGLDIA	2177	8	12	86
ALGLDIA	2178	8	12	86
ALGLDIA	2179	8	12	86
ALGLDIA	2180	8	12	86
ALGLDIA	2181	8	12	86
ALGLDIA	2182	8	12	86
ALGLDIA	2183			

HCY B62 Super Motif

[illegible]

HCY B62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy [%]
RRLMATHF	2875	8	12	86
RPTQNPFL	2259	8	11	79
RCMAGGDN	2243	8	12	86
RVCEKRAL	2821	8	100	100
RVESSNIV	2252	8	14	86
RACDTHIV	1755	8	11	79
RRLMATHF	175	8	14	100
RRLMATHF	1470	8	14	100
SPQENIV	2331	8	11	79
SPQENIV	2849	8	11	79
SPQENIV	2209	8	13	83
SVATLGF	1262	8	14	100
TIMAKNEV	2590	8	11	79
TLGEGATM	1622	8	13	83
TLGEGATM	1785	8	11	79
TLGEGATM	1871	8	11	79
TLGEGATM	1875	8	11	79
TLGEGATM	223	8	12	86
TLGEGATM	1464	8	12	86
TLGEGATM	1455	8	12	79
TLGEGATM	1871	8	13	86
TLGEGATM	1521	8	12	86
TLGEGATM	1337	8	13	93
TLGEGATM	157	8	14	100
TLGEGATM	1268	8	12	86
TLGEGATM	1686	8	14	86
TLGEGATM	1755	8	14	100
TLGEGATM	2639	8	11	78
TLGEGATM	1840	8	12	88
TLGEGATM	1918	8	14	100
TLGEGATM	1439	8	11	79
TLGEGATM	1898	8	11	79
TLGEGATM	1680	8	12	86
TLGEGATM	1920	8	14	86
TLGEGATM	1755	8	13	86
TLGEGATM	1779	8	14	100
TLGEGATM	616	8	14	100
TLGEGATM	1929	8	12	86
TLGEGATM	1880	8	12	86
TLGEGATM	150	8	14	100
TLGEGATM	1898	8	12	86
TLGEGATM	1886	8	11	79

HCY B62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Consistency [%]
APPSWCOM	1604	9	12	86
AFLTWARM	2869	9	11	79
ADGTRALV	1251	9	11	79
ACQVWPL	77	9	12	86
ADVMARLI	1812	9	12	100
CHSDLEWV	1653	9	11	79
ELKASREV	279	9	11	79
QLENTSTW	1657	9	12	86
QUNGFYLV	132	9	11	79
DLWLLPAL	1883	9	11	79
DLVNCESA	2772	9	11	79
DLVLTTHA	1134	9	12	86
DPQLSDSW	2410	9	12	79
DPFRGPNL	1111	9	12	86
EPPTGGAI	1372	9	13	93
ESKSTWLV	2245	9	12	86
ENTSTWLV	1666	9	12	86
FRSDIOLA	1773	9	12	86
FLALLSCL	177	9	14	100
FLLLADRV	728	9	12	86
FOVSGQW	2646	9	13	93
GRFTLQDA	1532	9	12	86
GRVQVQAL	1532	9	13	93
GLRLANAV	956	9	14	100
QLTHDAHF	1569	9	13	93
GRGEGANOW	1912	9	12	86
GPTELVLV	1625	9	14	100
GGNGGYLV	28	9	12	86
GNAGGLVAF	1643	9	12	86
GLALALNAV	1693	9	12	86
GNVTRNL	181	9	11	79
GNVTRNL	2519	9	14	100
GNVLEQV	154	9	13	93
HLHNVQV	636	9	12	86
HLPIVDSGM	1718	9	11	79
HMVNFISQ	1769	9	13	93
HCNVDQV	698	9	11	79
HCNVDQV	1570	9	11	79
HCNVDQV	1459	9	11	79
ELSGALNV	1881	9	11	79
KAVLQVPSV	1259	9	13	93
UTSCSSNV	2815	9	14	100
LWPFQVGV	2812	9	14	100
LFLLQLDA	726	9	13	93
LFNLQGV	1612	9	12	86

HCV B62 Super Motif (No binding data)

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)
LIPRGGRL	36	8	13	83
LPAISPPA	1888	8	13	93
LPALSTGL	687	9	12	86
LPCEPDV	2165	8	12	88
LPCCFSF	163	9	13	93
LVGGMAAL	1867	9	12	86
LVNPSVAA	1257	9	14	100
LVNLPAL	1884	9	11	79
LVTRHADN	1127	9	11	79
LVTRHADN	1817	9	11	79
NLGGVMAA	1815	9	11	79
NRTGGVTH	1282	9	12	86
NYGVGVLY	700	9	12	86
NLGGVDTL	1118	8	12	86
NLPCCSFI	168	9	13	93
NYGGLGVW	1106	9	11	79
NLGGVMAA	1815	9	11	79
PLYLGVAV	1628	9	13	93
PPRPGGQAV	1665	9	11	79
PPVHGGCL	2317	9	11	79
PPGVDTGL	2807	9	11	79
PIGDHLIF	1554	9	12	86
PIWSNLNE	2857	9	14	100
PIVSTVLL	2210	8	13	93
QVSGVMAA	2210	9	11	79
QPGVDELI	2868	9	11	79
QPGVFWLY	78	9	12	86
QPGVFWLY	57	9	13	93
QPGVFWLY	1029	9	12	86
RLAPITAT	2875	9	12	86
RLMLRTHF	2875	9	12	86
RVCERMAV	2821	9	14	100
RVCERMAV	2821	9	12	86
RVLGSENV	155	9	12	86
RVLGSENV	155	9	12	86
SALTPISH	2178	9	14	100
SPGALVGV	1893	9	13	93
SPGENTVA	2931	9	11	79
SPOORVEL	2649	9	11	79
SPGSGPSN	99	9	11	79
SVDDNTGV	1455	9	12	86
THAKRSTP	2897	9	11	79
THAPSTGL	1882	9	11	79
TLPGSPAL	886	9	11	79
TLTGGPAOL	125	9	12	86
TLWARMELM	2871	9	11	79
TLLVRLGA	1627	9	13	93

[illegible]

HCY B62 Super Motif

Sequence	Position	Peptide No.	No. of Amino Acids	Sequence Frequency	Consistency (%)
HKNDVQVL	986		10	11	79
LAGYDAGV	1856		10	11	79
ELGWYVCL	1816		10	12	79
WVSRKSLV	2159		10	11	79
KYVAGLSL	1777		10	14	100
MFPLGWRV	2613		10	11	79
KFLHGPTFL	1820		10	11	79
KVIDLTGCF	121		10	12	86
KVLNLFQV	1255		10	14	100
LLPLNGWV	1812		10	12	86
LLPLSPGA	1827		10	13	93
LMGTFPLGA	133		10	11	79
LLPLSPGAL	1168		10	12	86
LPKGRPLGV	189		10	13	93
LPKGRPLGV	37		10	13	93
LPKGRPLGV	1553		10	12	86
LVATGATVCA	1591		10	12	86
LVGLAGYGA	1853		10	11	79
LVGNLJALA	1887		10	12	86
LVWVWCAN	1737		10	11	79
LVWVWCAN	2172		10	11	100
LVWVWCAN	168		10	13	93
NFSVATLQIF	1260		10	14	100
PITYSTYQKF	1295		10	11	79
PLGGAAARLA	143		10	11	79
POPEYDLEI	2807		10	11	79
POCCELEW	2854		10	12	86
PNSVWVWVLA	2857		10	12	100
PNSVWVWVLA	508		10	13	93
QUPCEPQV	2164		10	12	88
QFPGGGKPPA	2401		10	11	79
RLKLSATSL	2818		10	11	79
RLVFFPLGV	2811		10	11	79
RLVWVWVWVW	317		10	12	86
RLNEDQNYA	138		10	12	96
SLVSTNGEL	281		10	12	96
SLVSTNGEL	1951		10	12	86
SPQALNSGV	1893		10	11	79
SQLSAPSLKA	2206		10	11	79
SPFPRRPRK	56		10	13	93
SVYATLGFQA	1262		10	14	100
TLHGPTFLY	1822		10	11	79
TLHGPTFLY	1168		10	12	86
TLHGPTFLY	686		10	11	79
TLHGPTFLY	125		10	12	88

HCY B62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy [%]
TPCTGGSSDL	1126	10	11	79
TPLLVRLGAV	1827	10	13	95
TPWNSMGNH	2886	10	10	86
TVQFSLDPTF	1486	10	12	86
VDITLGGFA	1472	10	13	86
VDAANALCL	1671	10	12	86
VDCAFETGA	1337	10	12	86
VLFPSVAITL	1258	10	14	100
VLITSGGNTL	2737	10	11	79
VLVGGVLAAL	1666	10	12	86
VLVLVFWAA	1356	10	14	100
VMSSTGPGY	2639	10	11	79
VPSDDAARY	1810	10	11	86
VPRKALGAF	1811	10	12	86
VWVAVQAL	1698	10	14	100
WLVGGVLAAL	1665	10	11	79
YLGSSGGPL	1165	10	12	86
YLPHPRPPL	35	10	12	86
YLYTRHDVI	1716	10	13	79
YVGGSGPPL	2176	10	12	86
YVGGSGPPL	1898	11	11	79
YVGGSGPPL	1235	11	13	93
YVGGSGPPL	2849	11	11	79
YVGGSGPPL	1602	11	12	86
YVGGSGPPL	1188	11	11	79
YVGGSGPPL	1917	11	11	100
YVGGSGPPL	1855	11	11	79
YVGGSGPPL	1657	11	12	86
YVGGSGPPL	2617	11	13	92
YVGGSGPPL	132	11	11	79
YVGGSGPPL	1134	11	11	79
YVGGSGPPL	1339	11	12	86
YVGGSGPPL	21	11	12	86
YVGGSGPPL	1731	11	12	86
YVGGSGPPL	1303	11	14	100
YVGGSGPPL	24	11	11	79
YVGGSGPPL	2646	11	14	100
YVGGSGPPL	1778	11	11	79
YVGGSGPPL	1562	11	12	86
YVGGSGPPL	1782	11	13	93
YVGGSGPPL	1625	11	13	93
YVGGSGPPL	1870	11	12	86
YVGGSGPPL	2618	11	14	100

HCY H62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Consistency [%]
GVALEDAWY	154	11	12	86
I-HONNDVY	696	11	11	78
HNNRFSQY	1769	11	13	83
HNNVDVRY	898	11	11	79
HQRECAQW	1870	11	11	79
ILGQVYALIA	1310	11	12	86
ILSGALDIA	1311	11	12	85
ILSGALWRY	1891	11	13	93
KPNSLFFDL	2608	11	11	79
KPTLHPTLL	1620	11	11	79
KQKALGLTA	1734	11	11	79
KVQDLTGFA	121	11	12	86
KVLNPSWA	1824	11	12	85
LKSSSSWY	1824	11	14	100
LKSSSWRY	2815	11	14	100
LVPFLDWR	2612	11	11	79
LLRLADRV	726	11	13	93
LENLGGWA	1812	11	12	83
LLPALSPQAL	1887	11	13	83
LPKRPFLGW	57	11	11	79
LPKRPSPSW	2240	11	11	88
LPKRPSPSW	1888	11	12	86
LPALSTGLIL	687	11	12	86
LPKCSFFLL	169	11	13	83
LPKCDLLEW	1553	11	12	85
LVNPSWAAL	1237	11	14	100
LYTRHADWPN	1897	11	11	79
LYKQKQAL	1815	11	11	78
NLDQMAQDL	1815	11	12	86
NTRMEKAV	2249	11	12	86
NLPALSTGA	1886	11	13	93
NLPKCSFEL	188	11	13	93
NTSTYGNFL	1295	11	11	79
PLEGTEGDL	2463	11	13	93
PKGPTDTRCF	2463	11	13	93
PKNSKQNL	1606	11	11	79
PVYCTTSPW	2857	11	12	86
PVYCTTSPW	508	11	13	93
RVYGVSEHL	635	11	13	93
RCEMGNTIR	2243	11	12	86
RCEMGALYOV	2821	11	12	86
SFLTALSCL	1776	11	12	86
SKLTPSNIA	2176	11	14	100

HCV B62 Super Motif

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy [%]
SPTHVRESDA	1835	11	12	86
SLPCPEPDA	2163	11	13	86
SVATIGRAY	1282	11	14	100
TLGFAYASKA	1266	11	12	88
TLNLEGGW	1811	11	12	86
TPCTQSSQLY	1126	11	11	79
TPGLVQDHL	1550	11	13	93
TPNSVLGNI	2858	11	12	88
TYLQMETAGA	1236	11	12	86
YLCETVAGCA	1256	11	11	79
YLVGQVLA	1852	11	11	78
YLVGQVLAIA	1666	11	12	86
YCPGSGRPA	2600	11	11	79
YQNNRLIAPA	1818	11	14	100
YVCAMLRHV	1801	11	11	79
YWLVGQVLAJ	1685	11	12	88
YXSSSGEPFL	1145	11	12	86
YXVAVGQVLA	1596	11	12	86
YGAIVGQVLA	1594	11	11	79
YGLQGSRL	276	11	11	88
YVPESDAARV	1539	11	12	86

428

Table XV
HCV Δ01 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*0101
ASTGSPY	165	8	20	100	
DNVLSRKY	737	10	18	90	0.0091
FAMPTGGY	631	10	19	95	0.0680
GFAFPGGQY	630	11	19	95	
GRETLEY	140	8	15	75	
GYSJRMFY	579	9	17	85	
HLWKAGLY	149	10	20	100	0.1100
KDAFTSPY	653	10	19	95	0.0001
LDITAGALY	30	9	17	95	0.0001
LDVDSMAY	115	11	17	95	0.0001
LDVDSMAY	117	11	15	75	0.0180
MAWYNGSY	380	11	17	85	0.0910
MSITLLEY	103	9	15	75	0.8500
NSVLSRKY	738	9	10	90	0.0005
PLDKGRPY	124	9	20	100	
PLDKGRPY	124	10	20	100	0.1700
PTTGRTSLY	797	9	17	85	0.2100
SASFGSPY	165	9	20	95	
SLVDSMAY	115	9	15	85	5.2000
SLVDSMAY	117	8	15	75	
TLTGRTSLY	798	8	17	85	
WLSLDSMAY	414	11	19	95	
WMAWYNGPS	359	11	17	85	0.3200
YPLMPLLY	640	8	19	95	
YSUNMGY	590	8	17	85	

25

Table XVI
HCY A03 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*0301
ACKWITIGER	647	10	12	86	0.0003
AARALAHVIR	147	10	11	79	
AATLFGGA	1264	8	14	100	
AATLFGAY	1264	9	14	100	
AVCTRGVA	1187	9	11	79	
AVCTRGVA	1187	10	11	79	
AVCTRGVAK	1187	11	11	79	
ADNVTIGER	648	9	12	86	
ADGCSGGA	1306	9	11	79	
ADGCSGGAY	1306	10	11	79	
ADVPVRR	1142	8	12	86	0.0003
ADVPVRR	1142	9	11	79	
AFASRONH	1928	8	14	100	
AGALVAFK	1005	8	12	86	
AGALVYLA	1344	9	12	86	
AGALVYLA	1344	11	11	79	
AGLSTLPDPA	1781	11	11	100	
AGVAGALVA	1882	9	12	86	
AGVAGALVF	1062	10	12	86	
AGVAGALVAFK	1852	11	12	86	
AGWLLSPR	94	8	12	86	0.0003
AGWLLSPRGR	94	11	12	86	
AGYAGVYA	1028	8	12	86	
AGYAGVYA	1028	10	12	86	
ALGLQTA	1727	8	12	86	
ALSTGLIH	609	8	12	86	
ALSTGLIH	609	10	12	86	
ALVGVVCA	1098	8	11	79	
ALVGVVCA	1888	10	11	79	
ALVGVVCA	1783	8	11	79	
ASLMAPTA	2200	11	11	79	0.0220
ASLMAPTA	2200	11	11	79	
ASLMAPTA	2200	11	11	79	
ASFGHVSPTH	1928	11	12	86	
ASSASQSLA	2204	10	14	100	
ATGNPGCSF	165	10	13	93	
ATLGFAY	1265	8	14	100	
ATLGFAYMSK	1265	11	12	86	
ATRTKTSR	48	8	11	79	
ATVCAHQA	1598	9	11	79	
AVCTRGVAK	1188	8	11	79	0.0220
AVCTRGVAK	1188	9	11	79	
AVCTRGVAK	1188	10	11	79	
AVCTRGVAK	1188	11	11	79	
AVQWMMRLIA	1917	10	14	100	
AVQWMMRLIA	1917	11	14	100	
AVQWMMRLIA	1917	11	14	100	
CAALRH	1903	8	13	93	

HICV A03 Motif with Binding Information

[illegible]

ILCV A03 Motif with Binding Information

[illegible]

UICV A03 Motif with Binding Information

Sequence	Position	No. of Ambos Adds	Sequence Frequency	Conservancy (%)	A*0301
KQDEIAKK	1404	8	12	86	
KQDEIAKR	2553	10	12	86	
KQDEIAKQ	1301	8	11	79	
KQDEIAFCH	1301	10	11	79	
KQDEIPAR	2604	8	11	79	
KLGWPLAR	2644	8	12	86	
KSTKVPAA	1241	8	12	86	
KSTKVPAA	1241	9	12	86	0.0009
KSTKVPAA	1241	11	12	86	
KSTKVPAA	1241	11	11	79	
KTRKNTNR	10	8	12	86	
KTRKNTNR	10	9	12	86	0.0110
KTRKRDPR	51	11	12	92	0.1600
KTRKRDPR	51	11	12	86	
KVDLITGE	121	10	12	86	
KVDLITGFA	121	11	12	86	
KVUNPSPA	1255	10	14	100	
KVUNPSPA	1255	11	14	100	
KVPAAYAA	1244	8	11	79	
KVPAAYAA	1305	10	11	79	
LAQGGSSGGA	1305	8	11	79	
LAQGGSSGGA	1729	8	12	86	
LAQGGSSGGA	1729	9	12	86	
LAQGGSSGGA	1657	9	11	79	
LAQGGSSGGA	1057	11	11	79	
LAQGGSSGGA	1552	10	11	79	
LDQNETAGA	1330	9	12	86	
LDQNETAGA	1330	10	12	86	
LELLIADA	727	8	14	100	
LELLIADA	727	9	12	86	
LEMLGGHWA	1813	11	12	86	
LEMLGGHWA	1813	11	12	86	
LEFEPBRI	280	9	11	79	
LGFGAYNSK	1287	8	12	88	0.0810
LGFGAYNSK	1287	10	12	86	
LGFGAYNSK	1257	11	12	86	
LGGGAHALA	144	9	11	79	
LGGGAHALA	144	10	11	79	
LGGWVAOLA	1817	10	12	86	
LGGWVAOLA	1332	10	13	86	
LGVPRATK	44	8	12	86	
LGVPRATK	2618	8	12	100	
LGVPRATK	1824	10	14	100	
LGVPRATK	2235	9	12	88	0.0008

UCV A03 Motif with Binding Information

[illegible]

HCY A03 Motif with Binding Information

[illegible]

HCV A03 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*0301
RAEDNNY	156	9	1174.17	86	0.0120
RALEDGNYA	156	10		86	
SAFLSHY	2923	8		73	
SAQSLSAPSLK	2267	11		73	
SCSSNSVA	2818	9		100	
SCSSNSVAH	2818	10		86	
SLYLVR	1133	8		86	
SLYLVRH	1133	9		86	
SLYLVRH	1133	10		86	
SLYLVRH	1133	11		86	
SLYLVRH	1133	12		86	
SLYLVRH	1133	13		86	
SLYLVRH	1133	14		100	
SLYLVRH	1133	15		100	
SLYLVRH	1133	16		100	
SLYLVRH	1133	17		100	
SLYLVRH	1133	18		100	
SLYLVRH	1133	19		100	
SLYLVRH	1133	20		100	
SLYLVRH	1133	21		100	
SLYLVRH	1133	22		100	
SLYLVRH	1133	23		100	
SLYLVRH	1133	24		100	
SLYLVRH	1133	25		100	
SLYLVRH	1133	26		100	
SLYLVRH	1133	27		100	
SLYLVRH	1133	28		100	
SLYLVRH	1133	29		100	
SLYLVRH	1133	30		100	
SLYLVRH	1133	31		100	
SLYLVRH	1133	32		100	
SLYLVRH	1133	33		100	
SLYLVRH	1133	34		100	
SLYLVRH	1133	35		100	
SLYLVRH	1133	36		100	
SLYLVRH	1133	37		100	
SLYLVRH	1133	38		100	
SLYLVRH	1133	39		100	
SLYLVRH	1133	40		100	
SLYLVRH	1133	41		100	
SLYLVRH	1133	42		100	
SLYLVRH	1133	43		100	
SLYLVRH	1133	44		100	
SLYLVRH	1133	45		100	
SLYLVRH	1133	46		100	
SLYLVRH	1133	47		100	
SLYLVRH	1133	48		100	
SLYLVRH	1133	49		100	
SLYLVRH	1133	50		100	
SLYLVRH	1133	51		100	
SLYLVRH	1133	52		100	
SLYLVRH	1133	53		100	
SLYLVRH	1133	54		100	
SLYLVRH	1133	55		100	
SLYLVRH	1133	56		100	
SLYLVRH	1133	57		100	
SLYLVRH	1133	58		100	
SLYLVRH	1133	59		100	
SLYLVRH	1133	60		100	
SLYLVRH	1133	61		100	
SLYLVRH	1133	62		100	
SLYLVRH	1133	63		100	
SLYLVRH	1133	64		100	
SLYLVRH	1133	65		100	
SLYLVRH	1133	66		100	
SLYLVRH	1133	67		100	
SLYLVRH	1133	68		100	
SLYLVRH	1133	69		100	
SLYLVRH	1133	70		100	
SLYLVRH	1133	71		100	
SLYLVRH	1133	72		100	
SLYLVRH	1133	73		100	
SLYLVRH	1133	74		100	
SLYLVRH	1133	75		100	
SLYLVRH	1133	76		100	
SLYLVRH	1133	77		100	
SLYLVRH	1133	78		100	
SLYLVRH	1133	79		100	
SLYLVRH	1133	80		100	
SLYLVRH	1133	81		100	
SLYLVRH	1133	82		100	
SLYLVRH	1133	83		100	
SLYLVRH	1133	84		100	
SLYLVRH	1133	85		100	
SLYLVRH	1133	86		100	
SLYLVRH	1133	87		100	
SLYLVRH	1133	88		100	
SLYLVRH	1133	89		100	
SLYLVRH	1133	90		100	
SLYLVRH	1133	91		100	
SLYLVRH	1133	92		100	
SLYLVRH	1133	93		100	
SLYLVRH	1133	94		100	
SLYLVRH	1133	95		100	
SLYLVRH	1133	96		100	
SLYLVRH	1133	97		100	
SLYLVRH	1133	98		100	
SLYLVRH	1133	99		100	
SLYLVRH	1133	100		100	
SLYLVRH	1133	101		100	
SLYLVRH	1133	102		100	
SLYLVRH	1133	103		100	
SLYLVRH	1133	104		100	
SLYLVRH	1133	105		100	
SLYLVRH	1133	106		100	
SLYLVRH	1133	107		100	
SLYLVRH	1133	108		100	
SLYLVRH	1133	109		100	
SLYLVRH	1133	110		100	
SLYLVRH	1133	111		100	
SLYLVRH	1133	112		100	
SLYLVRH	1133	113		100	
SLYLVRH	1133	114		100	
SLYLVRH	1133	115		100	
SLYLVRH	1133	116		100	
SLYLVRH	1133	117		100	
SLYLVRH	1133	118		100	
SLYLVRH	1133	119		100	
SLYLVRH	1133	120		100	
SLYLVRH	1133	121		100	
SLYLVRH	1133	122		100	
SLYLVRH	1133	123		100	
SLYLVRH	1133	124		100	
SLYLVRH	1133	125		100	
SLYLVRH	1133	126		100	
SLYLVRH	1133	127		100	
SLYLVRH	1133	128		100	
SLYLVRH	1133	129		100	
SLYLVRH	1133	130		100	
SLYLVRH	1133	131		100	
SLYLVRH	1133	132		100	
SLYLVRH	1133	133		100	
SLYLVRH	1133	134		100	
SLYLVRH	1133	135		100	
SLYLVRH	1133	136		100	
SLYLVRH	1133	137		100	
SLYLVRH	1133	138		100	
SLYLVRH	1133	139		100	
SLYLVRH	1133	140		100	
SLYLVRH	1133	141		100	
SLYLVRH	1133	142		100	
SLYLVRH	1133	143		100	
SLYLVRH	1133	144		100	
SLYLVRH	1133	145		100	
SLYLVRH	1133	146		100	
SLYLVRH	1133	147		100	
SLYLVRH	1133	148		100	
SLYLVRH	1133	149		100	
SLYLVRH	1133	150		100	
SLYLVRH	1133	151		100	
SLYLVRH	1133	152		100	
SLYLVRH	1133	153		100	
SLYLVRH	1133	154		100	
SLYLVRH	1133	155		100	
SLYLVRH	1133	156		100	
SLYLVRH	1133	157		100	
SLYLVRH	1133	158		100	
SLYLVRH	1133	159		100	
SLYLVRH	1133	160		100	
SLYLVRH	1133	161		100	
SLYLVRH	1133	162		100	
SLYLVRH	1133	163		100	
SLYLVRH	1133	164		100	
SLYLVRH	1133	165		100	
SLYLVRH	1133	166		100	
SLYLVRH	1133	167		100	
SLYLVRH	1133	168		100	
SLYLVRH	1133	169		100	
SLYLVRH	1133	170		100	
SLYLVRH	1133	171		100	
SLYLVRH	1133	172		100	
SLYLVRH	1133	173		100	
SLYLVRH	1133	174		100	
SLYLVRH	1133	175		100	
SLYLVRH	1133	176		100	
SLYLVRH	1133	177		100	
SLYLVRH	1133	178		100	
SLYLVRH	1133	179		100	
SLYLVRH	1133	180		100	
SLYLVRH	1133	181		100	
SLYLVRH	1133	182		100	
SLYLVRH	1133	183		100	
SLYLVRH	1133	184		100	
SLYLVRH	1133	185		100	
SLYLVRH	1133	186		100	
SLYLVRH	1133	187		100	
SLYLVRH	1133	188		100	
SLYLVRH	1133	189		100	
SLYLVRH	1133	190		100	
SLYLVRH	1133	191		100	
SLYLVRH	1133	192		100	
SLYLVRH	1133	193		100	
SLYLVRH	1133	194		100	
SLYLVRH	1133	195		100	
SLYLVRH	1133	196		100	
SLYLVRH	1133	197		100	
SLYLVRH	1133	198		100	
SLYLVRH	1133	199		100	
SLYLVRH	1133	200		100	
SLYLVRH	1133	201		100	
SLYLVRH	1133	202		100	
SLYLVRH	1133	203		100	
SLYLVRH	1133	204		100	
SLYLVRH	1133	205		100	
SLYLVRH	1133	206		100	
SLYLVRH	1133	207		100	
SLYLVRH	1133	208		100	
SLYLVRH	1133	209		100	
SLYLVRH	1133	210		100	
SLYLVRH	1133	211		100	
SLYLVRH	1133	212		100	
SLYLVRH	1133	213		100	
SLYLVRH	1133	214		100	
SLYLVRH	1133	215		100	
SLYLVRH	1133	216		100	
SLYLVRH	1133	217		100	
SLYLVRH	1133	218		100	
SLYLVRH	1133	219		100	
SLYLVRH	1133	220		100	
SLYLVRH	1133	221		100	
SLYLVRH	1133	222		100	
SLYLVRH	1133	223		100	
SLYLVRH	1133	224		100	
SLYLVRH	1133	225		100	
SLYLVRH	1133	226		100	
SLYLVRH	1133	227		100	
SLYLVRH	1133	228		100	
SLYLVRH	1133	229		100	
SLYLVRH	1133	230		100	
SLYLVRH	1133	231		100	
SLYLVRH	1133	232		100	
SLYLVRH	1133	233		100	
SLYLVRH	1133	234		100	
SLYLVRH	1133	235		100	
SLYLVRH	1133	236		100	
SLYLVRH	1133	237		100	
SLYLVRH	1133	238		100	
SLYLVRH	1133	239		100	
SLYLVRH	1133	240		100	
SLYLVRH	1133	241		100	
SLYLVRH	1133	242		100	
SLYLVRH	1133	243		100	
SLYLVRH	1133	244		100	
SLYLVRH	1133	245		100	
SLYLVRH	1133	246		100	
SLYLVRH	1133	247		100	
SLYLVRH	1133	248		100	
SLYLVRH	1133	249		100	
SLYLVRH	1133	250		100	
SLYLVRH	1133	251		100	
SLYLVRH	1133	252		100	
SLYLVRH	1133	253		100	
SLYLVRH	1133	254		100	
SLYLVRH	1133	255		100	
SLYLVRH	1133	256		100	
SLYLVRH	1133	257		100	
SLYLVRH	1133	258		100	
SLYLVRH	1133	259		100	
SLYLVRH	1133	260		100	
SLYLVRH	1133	261		100	
SLYLVRH	1133	262		100	
SLYLVRH	1133	263		100	
SLYLVRH	1133	264		100	
SLYLVRH	1133	265		100	
SLYLVRH	1133	266		100	
SLYLVRH	1133	267		100	
SLYLVRH	1133	268		100	
SLYLVRH	1133	269		100	
SLYLVRH	1133	270		100	
SLYLVRH	1133	271		100	

ICV A03 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*0301
TGNLPGCSF	166	9	13	93	
TSGSGKSTK	1237	9	12	85	
TSGSGKSTKPA	1237	9	12	86	
TLHAKGKPK	1266	9	11	78	
TLGFGANKK	1266	10	12	86	0.0810
TLGFGANKSKA	1266	11	12	86	
TLHGPTLLY	1622	10	11	79	0.0390
TUHGPTFLYR	1622	11	11	79	
TLPALSTGLH	666	11	11	79	
TLWARMALMTH	2071	11	11	79	
TSCSSNVSVA	2817	10	14	100	
TSEFSDPR	52	11	12	86	
TSEFSDPRK	52	11	12	86	
TSEFSDPRKGR	52	11	12	86	0.0003
TSLTGDK	1050	8	12	86	
TSMIDPSH	2177	9	13	93	0.0003
TTIMAGNEVF	2589	10	11	79	
TMHSPVF	1208	8	12	86	
TGARGGA	1597	8	11	78	
TGSLDPTF	1418	10	12	86	
TGSLDPTFGR	1338	10	12	86	
TALDQNTAGA	1338	9	14	100	
VAATLGFQA	1263	11	12	86	
VAATLGFQAY	1263	9	14	100	
VAGALVAF	1864	10	14	100	
VAGALVAFK	1864	8	12	86	
VAGALVAFK	1864	9	12	86	0.2400
VAGALVAFK	1592	12	12	86	
VAGATVCA	1592	10	11	79	0.0005
VAGATVCAK	1592	11	11	79	
VACALLER	1902	8	11	79	
VICALLERH	1902	9	11	79	
VCEKALY	2622	8	14	100	
VOSPYCF	505	8	13	93	
VODLLEF	1556	8	12	86	
VICTRGVAK	1189	9	11	78	
VICTRGVAKA	1189	9	11	78	
VONTYTHGA	1682	9	11	78	
VONTYTHGA	1682	9	14	100	
VOLADYCA	1654	9	11	79	
VODPYLWH	614	9	13	93	
VODPYLWH	614	10	13	93	
VFOVDEK	2597	8	12	86	
VFOVDEKGR	2597	11	11	79	
VFDLGR	2614	8	11	79	

ICV A03 Motif with Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*0301
VFTGLTHDA	1568	10	13	93	
VFTGLTHDAH	1566	11	13	93	
MGXGGSIF	277	9	12	86	
VGGVLAALA	1668	9	12	86	
VGGVLAALAA	1668	10	12	86	
VGGVLAALAA	1668	11	12	86	
VGGVLLPFI	31	9	13	93	0.0083
VGKVLPR	31	10	13	93	
VGKVLPR	316	9	13	93	
VGKVLPR	1009	10	11	79	0.0007
VGKVLPR	1009	11	11	79	
VGVWCAILR	122	9	12	86	
VGVWCAILR	122	10	12	86	
VGVWCAILR	1671	10	12	86	
VGVWCAILR	1671	11	12	86	
VGVWCAILR	1671	12	12	86	
VGVWCAILR	1671	13	13	93	
VGVWCAILR	1671	14	14	100	
VGVWCAILR	1671	15	15	100	
VGVWCAILR	1671	16	16	100	
VGVWCAILR	1671	17	17	100	
VGVWCAILR	1671	18	18	100	
VGVWCAILR	1671	19	19	100	
VGVWCAILR	1671	20	20	100	
VGVWCAILR	1671	21	21	100	
VGVWCAILR	1671	22	22	100	
VGVWCAILR	1671	23	23	100	
VGVWCAILR	1671	24	24	100	
VGVWCAILR	1671	25	25	100	
VGVWCAILR	1671	26	26	100	
VGVWCAILR	1671	27	27	100	
VGVWCAILR	1671	28	28	100	
VGVWCAILR	1671	29	29	100	
VGVWCAILR	1671	30	30	100	
VGVWCAILR	1671	31	31	100	
VGVWCAILR	1671	32	32	100	
VGVWCAILR	1671	33	33	100	
VGVWCAILR	1671	34	34	100	
VGVWCAILR	1671	35	35	100	
VGVWCAILR	1671	36	36	100	
VGVWCAILR	1671	37	37	100	
VGVWCAILR	1671	38	38	100	
VGVWCAILR	1671	39	39	100	
VGVWCAILR	1671	40	40	100	
VGVWCAILR	1671	41	41	100	
VGVWCAILR	1671	42	42	100	
VGVWCAILR	1671	43	43	100	
VGVWCAILR	1671	44	44	100	
VGVWCAILR	1671	45	45	100	
VGVWCAILR	1671	46	46	100	
VGVWCAILR	1671	47	47	100	
VGVWCAILR	1671	48	48	100	
VGVWCAILR	1671	49	49	100	
VGVWCAILR	1671	50	50	100	
VGVWCAILR	1671	51	51	100	
VGVWCAILR	1671	52	52	100	
VGVWCAILR	1671	53	53	100	
VGVWCAILR	1671	54	54	100	
VGVWCAILR	1671	55	55	100	
VGVWCAILR	1671	56	56	100	
VGVWCAILR	1671	57	57	100	
VGVWCAILR	1671	58	58	100	
VGVWCAILR	1671	59	59	100	
VGVWCAILR	1671	60	60	100	
VGVWCAILR	1671	61	61	100	
VGVWCAILR	1671	62	62	100	
VGVWCAILR	1671	63	63	100	
VGVWCAILR	1671	64	64	100	
VGVWCAILR	1671	65	65	100	
VGVWCAILR	1671	66	66	100	
VGVWCAILR	1671	67	67	100	
VGVWCAILR	1671	68	68	100	
VGVWCAILR	1671	69	69	100	
VGVWCAILR	1671	70	70	100	
VGVWCAILR	1671	71	71	100	
VGVWCAILR	1671	72	72	100	
VGVWCAILR	1671	73	73	100	
VGVWCAILR	1671	74	74	100	
VGVWCAILR	1671	75	75	100	
VGVWCAILR	1671	76	76	100	
VGVWCAILR	1671	77	77	100	
VGVWCAILR	1671	78	78	100	
VGVWCAILR	1671	79	79	100	
VGVWCAILR	1671	80	80	100	
VGVWCAILR	1671	81	81	100	
VGVWCAILR	1671	82	82	100	
VGVWCAILR	1671	83	83	100	
VGVWCAILR	1671	84	84	100	
VGVWCAILR	1671	85	85	100	
VGVWCAILR	1671	86	86	100	
VGVWCAILR	1671	87	87	100	
VGVWCAILR	1671	88	88	100	
VGVWCAILR	1671	89	89	100	
VGVWCAILR	1671	90	90	100	
VGVWCAILR	1671	91	91	100	
VGVWCAILR	1671	92	92	100	
VGVWCAILR	1671	93	93	100	
VGVWCAILR	1671	94	94	100	
VGVWCAILR	1671	95	95	100	
VGVWCAILR	1671	96	96	100	
VGVWCAILR	1671	97	97	100	
VGVWCAILR	1671	98	98	100	
VGVWCAILR	1671	99	99	100	
VGVWCAILR	1671	100	100	100	
VGVWCAILR	1671	101	101	100	
VGVWCAILR	1671	102	102	100	
VGVWCAILR	1671	103	103	100	
VGVWCAILR	1671	104	104	100	
VGVWCAILR	1671	105	105	100	
VGVWCAILR	1671	106	106	100	
VGVWCAILR	1671	107	107	100	
VGVWCAILR	1671	108	108	100	
VGVWCAILR	1671	109	109	100	
VGVWCAILR	1671	110	110	100	
VGVWCAILR	1671	111	111	100	
VGVWCAILR	1671	112	112	100	
VGVWCAILR	1671	113	113	100	
VGVWCAILR	1671	114	114	100	
VGVWCAILR	1671	115	115	100	
VGVWCAILR	1671	116	116	100	
VGVWCAILR	1671	117	117	100	
VGVWCAILR	1671	118	118	100	
VGVWCAILR	1671	119	119	100	
VGVWCAILR	1671	120	120	100	
VGVWCAILR	1671	121	121	100	
VGVWCAILR	1671	122	122	100	
VGVWCAILR	1671	123	123	100	
VGVWCAILR	1671	124	124	100	
VGVWCAILR	1671	125	125	100	
VGVWCAILR	1671	126	126	100	
VGVWCAILR	1671	127	127	100	
VGVWCAILR	1671	128	128	100	
VGVWCAILR	1671	129	129	100	
VGVWCAILR	1671	130	130	100	
VGVWCAILR	1671	131	131	100	
VGVWCAILR	1671	132	132	100	
VGVWCAILR	1671	133	133	100	
VGVWCAILR	1671	134	134	100	
VGVWCAILR	1671	135	135	100	
VGVWCAILR	1671	136	136	100	
VGVWCAILR	1671	137	137	100	
VGVWCAILR	1671	138	138	100	
VGVWCAILR	1671	139	139	100	
VGVWCAILR	1671	140	140	100	
VGVWCAILR	1671	141	141	100	
VGVWCAILR	1671	142	142	100	
VGVWCAILR	1671	143	143	100	
VGVWCAILR	1671	144	144	100	
VGVWCAILR	1671	145	145	100	
VGVWCAILR	1671	146	146	100	
VGVWCAILR	1671	147	147	100	
VGVWCAILR	1671	148	148	100	
VGVWCAILR	1671	149	149	100	
VGVWCAILR	1671	150	150	100	
VGVWCAILR	1671	151	151	100	
VGVWCAILR	1671	152	152	100	
VGVWCAILR	1671	153	153	100	
VGVWCAILR	1671	154	154	100	
VGVWCAILR	1671	155	155	100	
VGVWCAILR	1671	156	156	100	
VGVWCAILR	1671	157	157	100	
VGVWCAILR	1671	158	158	100	
VGVWCAILR	1671	159	159	100	
VGVWCAILR	1671	160	160	100	
VGVWCAILR	1671	161	161	100	
VGVWCAILR	1671	162	162	100	
VGVWCAILR	1671	163	163	100	
VGVWCAILR	1671	164	164	100	
VGVWCAILR	1671	165	165	100	
VGVWCAILR	1671	166	166	100	
VGVWCAILR	1671	167	167	100	
VGVWCAILR	1671	168	168	100	
VGVWCAILR	1671	169	169	100	
VGVWCAILR	1671	170	170	100	
VGVWCAILR	1671	171	171	100	
VGVWCAILR	1671	172	172	100	
VGVWCAILR	1671	173	173	100	
VGVWCAILR	1671	174	174	100	
VGVWCAILR	1671	175	175	100	
VGVWCAILR	1671	176	176	100	
VGVWCAILR	1671	177	177	100	
VGVWCAILR	1671	178	178	100	
VGVWCAILR	1671	179	179	100	
VGVWCAILR	1671	180	180	100	
VGVWCAILR	1671	181	181	100	
VGVWCAILR	1671	182	182	100	
VGVWCAILR	1671	183	183	100	
VGVWCAILR	1671	184	184	100	
VGVWCAILR	1671	185	185	100	
VGVWCAILR	1671	186	186	100	
VGVWCAILR	1671	187	187	100	
VGVWCAILR	1671	188	188	100	
VGVWCAILR	1671	189	189	100	
VGVWCAILR	1671	190	190	100	
VGVWCAILR	1671	191	191	100	
VGVWCAILR	1671	192	192	100	
VGVWCAILR	1671	193	193	100	
VGVWCAILR	1671	194	194	100	
VGVWCAILR	1671	195	195	100	
VGVWCAILR	1671	196	196	100	
VGVWCAILR	1671	197	197	100	
VGVWCAILR	1671	198	198	100	
VGVWCAILR	1671	199	199	100	
VGVWCAILR	1671	200	200	100	
VGVWCAILR	1671	201	201	100	
VGVWCAILR	1671	202	202	100	
VGVWCAILR	1671	203	203	100	
VGVWCAILR	1671	204	204	100	
VGVWCAILR	1671	205	205	100	
VGVWCAILR	1671	206	206	100	
VGVWCAILR	1671	207	207	100	
VGVWCAILR	1671	208	208	100	
VGVWCAILR	1671	209	209	100	
VGVWCAILR	1671	210	210	100	
VGVWCAILR	1671	211	211	100	
VGVWCAILR	1671	212	212	100	
VGVWCAILR	1671	213	213	100	
VGVWCAILR	1671	214	214	100	
VGVWCAILR	1671	215	215	100	
VGVWCAILR	1671	216	216	100	
VGVWCAILR	1671	217	217	100	
VGVWCAILR	1671	218	218	100	
VGVWCAILR	1671	219	219	100	
VGVWCAILR	1671	220	220	100	
VGVWCAILR	1671	221	221	100	
VGVWCAILR	1671	222	222	100	
VGVWCAILR	1671	223	223	100	
VGVWCAILR	1671	224	224	100	
VGVWCAILR	1671	225	225	100	
VGVWCAILR	1671	226	226	100	
VGVWCAILR	1671	227	227	100	
VGVWCAILR	1671	228	228	100	
VGVWCAILR	1671	229	229	100	
VGVWCAILR	1671	230	230	100	
VGVWCAILR	1671	231	231	100	
VGVWCAILR	1671	232	232	100	
VGVWCAILR	1671	233	233	100	
VGVWCAILR	1671	234	234		

Table XVII

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*1101
ACNWTIRGR	647	10	12	88	
ACNCTGVAK	1167	10	11	79	
ATLFGAY	1264	9	14	100	
ACNCTGVAK	1107	10	11	79	
ACNWTIRGR	648	9	12	86	
ADQCGSGAY	1306	10	11	79	
ADNPVRR	1142	8	12	86	
ADNPVRR	1142	9	11	79	
AFASRQH	1926	8	14	100	
AGALVAK	1805	8	12	86	
AGVAGLVAK	1802	11	12	86	
ADKLSPR	94	8	12	86	
AGLSTGRSR	689	11	12	86	
ALSTGLH	689	8	12	86	
ASQSAFLK	2208	10	12	86	0.0027
ASGRANNSTH	1928	11	12	86	
ATLFGAY	1265	8	14	100	
ATLFGAYAK	1265	11	12	90	
ATKTSER	48	8	11	79	
ACTGVAK	1180	9	11	79	
CAKATGR	1383	9	13	93	
CEPQVQY	128	9	13	93	0.0250
CGNLTCT	2742	8	11	79	
CGSSALVTR	1130	11	11	79	
CLRLGVPLRL	2841	11	12	86	
CNCSYGH	304	9	11	79	
CNWTIRGR	649	8	12	86	
CSSNSYAH	2019	9	12	86	
CTCGSSALVTR	1130	9	11	79	
CTCGSSALVTR	555	11	11	79	0.0083
COPEGRGR	2569	9	11	79	0.0005
COPEGRGR	2569	10	11	79	0.0008
DAFLSOTK	1574	9	14	100	0.0005
DGCGSGAY	1307	9	11	79	
DNIQEQH	1316	9	12	86	0.0002
DLVAVCK	2617	9	13	93	
DLVLVTH	1134	8	13	96	
DNPVRR	1143	10	11	79	
ECTGCGWY	1324	10	14	100	0.0014
EGANQWHR	2815	8	12	86	
EGANQWHR	2815	9	12	86	0.0270
FOVPEGRGR	2598	10	11	79	
FOVPEGRGR	2598	11	11	79	

ICV All Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*1101
FGAYNSKAH	1269	9	12	88	
FGAGQAVR	2554	9	12	86	0.0005
FLLLADAR	728	8	14	100	
FTEAMTRY	2782	8	14	100	
FTGLTHDAH	1587	10	13	79	
GAHALAH	146	8	11	79	
GAHNSPTSR	1931	11	11	79	
GAHNSPTSR	1916	8	14	100	
GAHNSPTSR	1270	8	12	86	
GAHNSPTSR	1270	8	13	93	
GAHNSPTSR	1268	8	12	86	
GAHNSPTSR	1268	10	12	86	
GAHNSPTSR	2645	9	11	79	
GAHNSPTSR	145	9	11	79	
GAHNSPTSR	1308	8	11	79	
GAHNSPTSR	26	10	14	100	
GAHNSPTSR	26	10	14	100	
GAHNSPTSR	1392	9	14	100	0.0001
GAHNSPTSR	1392	11	14	100	
GAHNSPTSR	1669	10	12	86	
GAHNSPTSR	32	13	13	93	
GAHNSPTSR	3027	9	13	93	0.0010
GAHNSPTSR	1052	8	13	93	
GAHNSPTSR	1052	8	13	93	
GAHNSPTSR	2821	8	11	79	
GAHNSPTSR	2821	8	11	79	
GAHNSPTSR	2821	10	11	79	0.0005
GAHNSPTSR	1569	8	13	93	
GAHNSPTSR	1931	9	12	86	
GAHNSPTSR	1931	9	12	86	
GAHNSPTSR	2248	11	12	86	
GAHNSPTSR	1131	11	12	86	
GAHNSPTSR	1131	11	12	86	
GAHNSPTSR	2641	11	11	79	
GAHNSPTSR	2063	8	11	79	
GAHNSPTSR	1863	10	12	86	1.4000
GAHNSPTSR	1081	8	11	79	
GAHNSPTSR	2035	10	11	79	0.0140
GAHNSPTSR	1670	9	12	79	0.0110
GAHNSPTSR	45	11	14	100	
GAHNSPTSR	269	11	12	86	
GAHNSPTSR	1514	11	11	79	
GAHNSPTSR	1509	8	11	79	
GAHNSPTSR	1509	10	11	79	
GAHNSPTSR	1500	11	11	79	

[illegible]

HCV A11 Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy [%]	A*1101
PGCVPOAR	224	8	12	86	
PGEGAWMAN	1913	11	13	93	
PGGGANGGV	25	11	14	100	
RGLPVCOZH	1551	9	13	93	
PGYPMWLY	79	8	14	100	
PITYSTYCK	1295	9	11	79	
PLGGAAHNAH	143	11	11	79	
PGHSTH	287	9	13	93	
PHWEDVR	128	11	13	93	
PSPVAVGTDI	514	11	13	93	
PSWOCMAK	1807	8	11	79	
PTDCFMH	507	8	13	93	
PTDPRRSR	109	9	12	86	0.0005
PTGSGKSTK	1235	9	13	93	0.0001
PTLHGPTLLY	1821	11	11	79	
PWAGTIDR	516	9	13	93	0.0005
QAEYAGAR	1340	6	12	86	
QKQVQVLR	28	11	13	93	
QLETPSR	289	9	11	79	
QLFISPR	289	9	11	79	0.0330
QLSAFSLK	2210	8	11	79	
QNVNDVY	699	8	11	79	
QNVNDVLY	699	10	11	79	
RAVACTIGVAK	1168	11	11	79	
RAVACTIGV	1168	8	14	100	
RATKRSER	47	8	14	100	
RGNHSPTH	1930	9	12	86	0.0001
RGNHSPTH	1930	10	12	86	0.0001
RSPILGVR	40	8	13	93	
RSPILGVFAR	40	11	11	79	
RSPHOPRK	59	9	13	93	0.0017
RSLSLPR	1154	8	12	86	
RSLSLPR	1154	8	11	79	
RLGKATIK	43	11	11	79	
RMLSAFSLH	2918	11	11	79	0.0290
RLMAFSR	1923	8	14	100	
RLAFASRGNH	1923	11	14	100	
RLWFLGLVR	2611	11	11	79	0.0270
RLLPATAY	1029	9	12	86	
RLNYGGVGH	635	9	14	100	0.0200
RLNYGGVGH	635	10	14	100	
RLNYHPPQDK	13	11	11	79	
RLNYHPPQDK	13	11	11	79	
RIVERMAYL	2821	9	14	100	0.5000
RIVERMAYL	2821	9	14	100	0.0068
RLLEDVNY	155	9	12	86	

HCY All Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy [%]	A*1101
SAFELMSV	2823	8	11	79	
SGKASAKYK	11	11	11	79	
SGSSNSVAH	2818	10	12	86	
SGSLVTR	1133	8	12	86	
SGSLVTR	1133	9	12	86	
SGASTKVPAY	1239	11	12	86	
SMATDPFH	2178	8	14	100	
SNSLRPH	2400	8	12	86	
SSDLVTR	1132	9	12	86	
SSDLVTR	1132	10	12	86	0.0044
SNKVSVAH	2020	8	12	86	0.0013
STGLRKH	691	8	12	86	
STKPKRPH	1942	8	12	86	
STNPKPR	2	8	11	79	
STNPKPRK	2	9	11	79	
STNPKPRKTK	2	11	11	79	
SVAATLGFQAY	1262	11	14	100	
TGGFAQMGY	127	10	13	93	
TGSSSLY	1129	8	11	79	
TGPRPRR	110	8	12	86	
TGPRPRK	110	8	12	86	
TGPRPRK	1468	9	13	93	0.0001
TGSGKTK	1237	8	13	93	
TLGFGAYMSK	1268	10	12	86	0.0010
TLHGPTFLYR	1622	10	11	79	0.0007
TLHGPTFLYR	1622	11	11	79	
TLPALSTGUH	880	11	11	79	
TLWARMUMTH	2871	11	11	79	
THPKPRK	3	8	11	79	
THPKPRKTK	3	9	11	79	
THPKPRKTKR	3	11	11	79	
THPKPRDK	15	9	11	79	
TSKSSNSVAH	2817	11	12	86	
TSERSQPR	52	8	13	93	0.0001
TSERSQPR	52	10	12	86	
TSERSQPRH	52	11	12	86	
TSLSGRUK	1050	9	12	86	
TSLSGRUK	1263	9	13	93	0.0001
VAGALVAFK	1664	10	14	100	0.8900
VAGALVAFK	1664	9	12	86	0.0038
VAYCATVCAR	1592	10	11	79	
VCAALLR	1802	8	11	79	
VCAALLRRH	1802	9	11	79	
VCAERWLY	2622	8	14	100	
VCTROVAK	1189	8	11	79	

HCV All Motif With Binding Information

[illegible]

Table XVIII
ICV A24 Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*2401
AVQKMMWV	319	8	12	86	
AYAGGKVL	1248	10	11	79	0.0009
AYTRGLDSVI	1421	11	14	100	
CYDAGCAW	1525	8	11	79	
CYDAGCAWEL	1525	11	11	79	
DFSLPIIF	1488	8	14	100	
DFSLPIIFL	1488	10	15	100	
FWKGMNIF	1765	9	12	86	6.9000
FWKGMNIF	1765	10	12	86	
FWKGMNIF	1765	11	13	93	
FWKGMNIF	1765	12	13	93	
FWKGMNIF	1765	13	13	93	
FWKGMNIF	1765	14	11	79	
FWKGMNIF	1765	15	11	79	
FWKGMNIF	1765	16	11	79	
FWKGMNIF	1765	17	11	79	
FWKGMNIF	1765	18	11	79	
FWKGMNIF	1765	19	11	79	
FWKGMNIF	1765	20	11	79	
FWKGMNIF	1765	21	11	79	
FWKGMNIF	1765	22	11	79	
FWKGMNIF	1765	23	11	79	
FWKGMNIF	1765	24	11	79	
FWKGMNIF	1765	25	11	79	
FWKGMNIF	1765	26	11	79	
FWKGMNIF	1765	27	11	79	
FWKGMNIF	1765	28	11	79	
FWKGMNIF	1765	29	11	79	
FWKGMNIF	1765	30	11	79	
FWKGMNIF	1765	31	11	79	
FWKGMNIF	1765	32	11	79	
FWKGMNIF	1765	33	11	79	
FWKGMNIF	1765	34	11	79	
FWKGMNIF	1765	35	11	79	
FWKGMNIF	1765	36	11	79	
FWKGMNIF	1765	37	11	79	
FWKGMNIF	1765	38	11	79	
FWKGMNIF	1765	39	11	79	
FWKGMNIF	1765	40	11	79	
FWKGMNIF	1765	41	11	79	
FWKGMNIF	1765	42	11	79	
FWKGMNIF	1765	43	11	79	
FWKGMNIF	1765	44	11	79	
FWKGMNIF	1765	45	11	79	
FWKGMNIF	1765	46	11	79	
FWKGMNIF	1765	47	11	79	
FWKGMNIF	1765	48	11	79	
FWKGMNIF	1765	49	11	79	
FWKGMNIF	1765	50	11	79	
FWKGMNIF	1765	51	11	79	
FWKGMNIF	1765	52	11	79	
FWKGMNIF	1765	53	11	79	
FWKGMNIF	1765	54	11	79	
FWKGMNIF	1765	55	11	79	
FWKGMNIF	1765	56	11	79	
FWKGMNIF	1765	57	11	79	
FWKGMNIF	1765	58	11	79	
FWKGMNIF	1765	59	11	79	
FWKGMNIF	1765	60	11	79	
FWKGMNIF	1765	61	11	79	
FWKGMNIF	1765	62	11	79	
FWKGMNIF	1765	63	11	79	
FWKGMNIF	1765	64	11	79	
FWKGMNIF	1765	65	11	79	
FWKGMNIF	1765	66	11	79	
FWKGMNIF	1765	67	11	79	
FWKGMNIF	1765	68	11	79	
FWKGMNIF	1765	69	11	79	
FWKGMNIF	1765	70	11	79	
FWKGMNIF	1765	71	11	79	
FWKGMNIF	1765	72	11	79	
FWKGMNIF	1765	73	11	79	
FWKGMNIF	1765	74	11	79	
FWKGMNIF	1765	75	11	79	
FWKGMNIF	1765	76	11	79	
FWKGMNIF	1765	77	11	79	
FWKGMNIF	1765	78	11	79	
FWKGMNIF	1765	79	11	79	
FWKGMNIF	1765	80	11	79	
FWKGMNIF	1765	81	11	79	
FWKGMNIF	1765	82	11	79	
FWKGMNIF	1765	83	11	79	
FWKGMNIF	1765	84	11	79	
FWKGMNIF	1765	85	11	79	
FWKGMNIF	1765	86	11	79	
FWKGMNIF	1765	87	11	79	
FWKGMNIF	1765	88	11	79	
FWKGMNIF	1765	89	11	79	
FWKGMNIF	1765	90	11	79	
FWKGMNIF	1765	91	11	79	
FWKGMNIF	1765	92	11	79	
FWKGMNIF	1765	93	11	79	
FWKGMNIF	1765	94	11	79	
FWKGMNIF	1765	95	11	79	
FWKGMNIF	1765	96	11	79	
FWKGMNIF	1765	97	11	79	
FWKGMNIF	1765	98	11	79	
FWKGMNIF	1765	99	11	79	
FWKGMNIF	1765	100	11	79	

ICV A24 Motif With Binding Information

Sequence	Position	No. of Amino Acids	Sequence Frequency	Conservancy (%)	A*2401
TWALVGM	1654	9	12	86	
TYSTYGF	1297	8	13	93	
TYSTYKFL	1297	9	12	86	0.0220
VFGLTH	1566	8	13	93	
WGSSTGF	2639	11	11	79	
VLLPFRFL	34	8	13	93	0.0016
WVSTGFL	1566	8	14	100	
YRGLDSN	1422	10	14	100	
53		2			

Table XIX a

[illegible]

[illegible]

HCV DR-Super Motif Binding Data Not Included

Core Sequences	Core First	Core Consensus (%)	Emulseq Sequence	Position in HCV Polyprotein	Emulseq Frequency	Sequence Conservation (%)
NS5A-1	12	86	NS5A-1A-1033-11	1658	9	57
NS5A-2	13	83	NS5A-1A-1033-11	1347	9	64
NS5A-3	12	86	NS5A-1A-1033-11	1347	12	86
NS5A-4	12	86	NS5A-1A-1033-11	1347	154	100
NS5A-5	13	83	NS5A-1A-1033-11	1347	154	60
NS5A-6	13	83	NS5A-1A-1033-11	1347	2172	64
NS5A-7	11	79	NS5A-1A-1033-11	1347	2172	71
NS5A-8	11	79	NS5A-1A-1033-11	1347	2172	71
NS5A-9	12	86	NS5A-1A-1033-11	1347	2172	85
NS5A-10	12	86	NS5A-1A-1033-11	1347	2172	85
NS5A-11	14	100	NS5A-1A-1033-11	1253	14	100
NS5A-12	12	86	NS5A-1A-1033-11	1253	14	100
NS5A-13	12	86	NS5A-1A-1033-11	1253	14	100
NS5A-14	12	86	NS5A-1A-1033-11	1253	14	100
NS5A-15	11	79	NS5A-1A-1033-11	1253	7	55
NS5A-16	12	86	NS5A-1A-1033-11	1253	7	55
NS5A-17	12	86	NS5A-1A-1033-11	1253	7	55
NS5A-18	12	86	NS5A-1A-1033-11	1253	7	55
NS5A-19	11	79	NS5A-1A-1033-11	1253	10	71
NS5A-20	11	79	NS5A-1A-1033-11	1253	10	71
NS5A-21	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-22	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-23	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-24	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-25	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-26	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-27	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-28	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-29	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-30	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-31	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-32	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-33	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-34	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-35	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-36	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-37	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-38	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-39	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-40	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-41	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-42	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-43	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-44	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-45	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-46	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-47	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-48	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-49	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-50	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-51	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-52	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-53	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-54	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-55	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-56	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-57	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-58	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-59	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-60	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-61	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-62	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-63	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-64	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-65	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-66	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-67	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-68	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-69	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-70	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-71	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-72	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-73	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-74	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-75	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-76	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-77	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-78	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-79	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-80	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-81	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-82	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-83	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-84	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-85	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-86	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-87	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-88	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-89	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-90	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-91	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-92	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-93	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-94	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-95	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-96	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-97	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-98	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-99	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-100	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-101	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-102	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-103	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-104	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-105	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-106	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-107	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-108	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-109	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-110	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-111	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-112	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-113	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-114	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-115	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-116	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-117	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-118	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-119	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-120	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-121	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-122	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-123	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-124	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-125	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-126	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-127	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-128	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-129	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-130	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-131	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-132	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-133	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-134	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-135	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-136	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-137	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-138	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-139	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-140	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-141	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-142	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-143	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-144	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-145	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-146	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-147	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-148	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-149	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-150	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-151	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-152	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-153	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-154	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-155	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-156	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-157	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-158	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-159	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-160	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-161	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-162	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-163	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-164	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-165	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-166	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-167	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-168	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-169	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-170	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-171	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-172	12	86	NS5A-1A-1033-11	1253	10	86
NS5A-173	12	86				

Table XIXb. JICY DR Super Motif With Binding Data

[illegible]

HICY DR Super Motif With Binding Data

[illegible]

Table XXb HCV DR 3A Motif With Binding Information

Core Sequence	Example Sequence	DR3	DR1	DR2=DR1	DR4=DR2	DR5=15	DR6=11	DR7=12	DR8=19	DR7	DR1=2	DR9	DR1=3
FLVGGGSA	YQPLVGGGSGAN		0.0001		0.1690					0.0005			
LEGGP	YQPLSDLPFTETI	-0.0017											
LEGGP	YQPLSDLPFTETI	-0.0017											
LPCEPDTV	GSGLVCEPDTVNA		0.0000	0.0015	0.0014	0.0079	0.0000	0.0000	0.0017	0.0017		0.0030	
MAWVMANN	GSRAWVMANNST		0.0004		0.0740				-0.0003				
MAVPSIET	LTBALVPSIETAT												
MSGLVET	MAKAGLEVETIV	1.1600	0.0046	0.0047	0.0014	0.0006	0.0029	0.0003	0.0003	0.0003			
MSGLVET	MAKAGLEVETIV	0.0003											
VQPLGVIV	GLVQPLGVIV												
VFTMGSPV	RLVFTMGSPVCK												
VFTMGSPV	RSPVFTMGSPVAP												
VLECGTQV	DSSVLECGTQAGZW	-0.0017			0.0000				-0.0002				
VLEDDVNA	DVYVLEDDVNAIGN		0.0007										
VLDVNGY	LRVAVLDVNGYAG												
VLDVNGY	LRVAVLDVNGYAG												
VLELITFC	GVVAVLELITFCRN		0.0003		0.0004				-0.0002				
VSEPLULP	GACTVSEPLULPIL												
VMDGLQSY	SAMVMDGLQSYLV	-0.0017											
YVPSDAAA	PIIVVPSDAAAVT	0.0020											

19

Table XXc HCV 3E Motif

Core Sequence	Core Freq	Core Consistency (%)	Empty Sequence	Residues in HCV Region	Empty Sequence Frequency	Empty Sequence Consistency (%)
TCSSKCD	14	100	IMFDSKNCSEA	1395	14	100
PSDFKCD	11	79	PAFSDYTRKDSV	2657	11	79
LASDFKCA	12	86	GAGLSEPKALGL	1726	8	57
MSDFKFT	11	73	LEKPLLEPTLL	1816	10	71
MSDFKFT	11	88	LEKPLLEPTLL	1816	10	71
YLDFKDFV	12	88	SEKPLLEPTLL	1120	11	78
ASDFKDFV	11	79	SEKPLLEPTLL	1	1	78

TABLE XXI. Population coverage with combined HLA Supertypes

<u>HLA-SUPERTYPES</u>	<u>PHENOTYPIC FREQUENCY</u>					
	Caucasian	North American Black	Japanese	Chinese	Hispanic	Average
<u>a. Individual Supertypes</u>						
A2	45.8	39.0	42.4	45.9	43.0	43.2
A3	37.5	42.1	45.8	52.7	43.1	44.2
B7	38.6	52.7	48.8	35.5	47.1	44.7
A1	47.1	16.1	21.8	14.7	26.3	25.2
A24	23.9	38.9	58.6	40.1	38.3	40.0
B44	43.0	21.2	42.9	39.1	39.0	37.0
B27	28.4	26.1	13.3	13.9	35.3	23.4
B62	12.6	4.8	36.5	25.4	11.1	18.1
B58	10.0	25.1	1.6	9.0	5.9	10.3
<u>b. Combined Supertypes</u>						
A2, A3, B7	83.0	86.1	87.5	88.4	86.3	86.2
A2, A3, B7, A24, B44, A1	99.5	98.1	100.0	99.5	99.4	99.3
A2, A3, B7, A24, B44, A1, B27, B62, B58	99.9	99.6	100.0	99.8	99.9	99.8

SP 184189 v1

HIC ANALOGS

AA	Sequence	Fixed Nomen.	A1 Motif	A2 Super Motif	A3 Super Motif	A24 Motif	B7 Super Motif	I* Anchor Fixer
g	CWNGCWAV		N	Y	N	N	N	
	40							

Table XXIII. Immunogenicity of identified supermotif-bearing peptides

Supermotif	Peptide	Sequence	Protein	Position	Human ^a				Transgenic mice ^b		
					Immunogenicity				Frequency	Response	
					Barnabat, patients	Barnabat, contacts	Chisari	Pape			overall
A2	1073.05	LLFNILGGWV	NS4	1812	1/6	7/17	2/21	0/6	10/50	6/6	6.4 (1.7)
	1090.18	FLLLDARV	NS1/E2	728	2/6	7/17	1/21	0/6	10/50	5/6	9.5 (3.0)
	1013.02	YLVA'QATV	NS4	1590	1/6	4/17	1/21	0/6	6/50	5/6	8.5 (3.7)
	1090.22	RLVFPDGV	NS5	2578	2/6	5/17	0/21	0/6	7/50	0/6	-
	1013.1002	DLMGYPLV	Core	132	2/6	7/17	1/21	1/6	11/50	5/6	8.8 (2.6)
	24.0073	WMNRLIIFA	NS4	1920	1/6	3/17	2/21	1/6	7/50	0/6	-
	24.0075	VLVGGVLA	NS4	1666	1/6	6/17	3/21	1/6	11/50	0/6	-
	1174.08	HMWNFISGI	NS4	1769	3/6	3/17	2/21	0/6	8/50	6/6	6.4 (1.7)
	1073.06	ILAGYGAGV	NS4	1851	2/6	3/17	0/21	0/6	5/50	3/6	54.7 (3.3)
	1073.07	YLLPRGPRL	CORE	35	2/6	5/17	7/21	1/6	17/50	4/6	59.1 (7.2)
	24.0071	LLFLLADA	NS1/E2	726	2/6	9/17	0/21	0/6	11/50	0/6	-
	1.0119	YLVTRHADV	NS3	1131	6/6	10/17	0/21	1/6	17/50	0/6	-
A3	1.0952	KTSRSQPR	CORE	51	2/16	1/4	3/12	0/6	6/38	3/6	23.4 (1.3)
	1073.11	RLGVRATRK	CORE	43	4/16	1/4	7/12	1/6	13/38	3/6	42.2 (1.2)
	1.0955	QLFTSPRR	ENV	290	1/16	0/4	6/12	1/6	8/38	2/6	2.8 (1.1)
	1073.13	RMVYGVGEHR	NS1/E2	632	5/16	1/4	4/12	1/6	11/38	3/6	4.4 (1.1)
	1.0123	LIFCHSKK	NS3	1396	6/16	1/4	4/12	2/6	13/38	6/6	56.5 (1.7)
	1073.10	GVAGALVAFK	NS4	1863	3/16	0/4	6/12	2/6	11/38	1/6	7.1
	24.0090	VAGALVAFK	NS4	1864	4/16	1/4	6/12	0/4	11/38	0/6	-
	24.0086	TLGFGATMSK	NS3	1262	6/16		2/12	2/5	10/33	5	-
	1145.12	LPQCSFSF	CORE	169			2	3/10			-

Table XXIV. Human and murine MHC-peptide binding assays established using purified MHC molecules and gel filtration chromatography

A. Class I binding assays				Radiolabeled peptide		Notes
Species	Antigen	Allele	Cell line	Source	Sequence	
Human	A1	A*0101	Scitelin	Hu. J chain 102-110	YTAVPLVY	no NEN in P cocktail
	A2	A*0201	JY	HBVc 18-27 F6->Y	FLPSDYFVS	"
	A2	A*0202	P815 (transfected)	HBVc 18-27 F6->Y	FLPSDYFVS	"
	A2	A*0203	FUN	HBVc 18-27 F6->Y	FLPSDYFVS	"
	A2	A*0206	CLA	HBVc 18-27 F6->Y	FLPSDYFVS	"
	A2	A*0207	721.221 (transfected)	HBVc 18-27 F6->Y	FLPSDYFVS	"
	A3		GM3107	non-natural (A3CON1)	KVFFYALNK	"
	A11		BVR	non-natural (A3CON1)	KVFFYALNK	"
	A24	A*2402	KAS116	non-natural (A24CON1)	AVDNTNKF	"
	A31	A*3101	SPACH	non-natural (A3CON1)	KVFFYALNK	"
	A33	A*3301	LWAGS	non-natural (A3CON1)	KVFFYALNK	"
	A28/68	A*6801	CIK	HBVc 141-151 T7->Y	STLPETVYVRR	"
	A28/68	A*6802	AMAI	HBVc 141-151 T7->Y	FTQAGYVRL	"
	B7	B*0702	GM3107	HBVc pol 646-654 C4->A	APRTLYLL	"
Mouse	B7	B*0702	Scitelin	AL sign seq. 513 (L2->Y)	FLKDYQLL	"
	B8	B*0801	LQ2	ILVp 386-393 Y1->F, Q3->S	FLKDYQLL	"
	B27	B*2705	CIK, BVR	R 608	FLKDYQLL	"
	B35	B*3501	TISI	non-natural (B35CON2)	FFPKYAAAF	"
	B35	B*3502	EHM	non-natural (B35CON2)	FFPKYAAAF	"
	B35	B*3503	PITOUT	non-natural (B35CON2)	FFPKYAAAF	"
	B44	B*4403		EF-1 G6->Y	FFPKYAAAF	"
	B51		KAS116	non-natural (B35CON2)	AEMGKTSFY	"
	B53	B*5301	AMAI	non-natural (B35CON2)	FFPKYAAAF	"
	B54	B*5401	KT3	non-natural (B35CON2)	FFPKYAAAF	"
	Cw4	Cw*0401	CIK	non-natural (C4CON1)	QYDDAVYKL	"
	Cw6	Cw*0602	721.221 transfected	non-natural (C6CON1)	YHHDGNNL	"
	Cw7	Cw*0702	721.221 transfected	non-natural (C6CON1)	YHHDGNNL	"
	D ^a		EL4	Adenovirus E1A P7->Y	SGFSNTYPEI	"
	K ^b		EL4	VSV NP 52-59	RGYVFOQL	"
	D ^d		P815	HIV-IIIb ENV G4->Y	RGYVFAFTI	"
	K ^d		P815	non-natural (KdCON1)	KRPMKTYI	"
	L ^d		P815	HBVc 28-39	IPQSLDSYWTSL	"

Table XXIV. Human and murine MHC-peptide binding assays established using purified MHC molecules and gel filtration chromatography

B. Class II binding assays		Radiolabelled peptide			Notes
Species	Antigen	Allele	Cell line	Source	
Human	DR1	DRB1*0101	L62	HA Y307-319	YFKYVKQNTLKLAT
	DR2	DRB1*1501	L466.1	MBP 88-102Y	VVHFKNIVTRTPPY
	DR3	DRB1*1601	L242.5	non-natural (760.16)	YAFAAAKTAFAA
	DR4 ^a	DRB1*0301	MAT	MT 65SD Y3-13	YKTIADFEARR
	DR4 ^a 10	DRB1*0401	Praxis	non-natural (717.01)	YARFQSTTLKQKT
	DR4 ^a 14	DRB1*0402	YAR	non-natural (717.10)	YARFQSTTLK&AAA
	DR4 ^a 15	DRB1*0404	BRN-40	non-natural (717.01)	YARFQSTTLKQKT
	DR7	DRB1*0405	KT3	non-natural (717.01)	YARFQSTTLKQKT
	DR8	DRB1*0701	Pfurot	Tet. tox. 830-845	QYICAKNSKEGITE
	DR8	DRB1*0802	OLL	Tet. tox. 830-845	QYICAKNSKEGITE
	DR9	DRB1*0901	HID	Tet. tox. 830-845	QYICAKNSKEGITE
	DR11	DRB1*1101	Sweig	Tet. tox. 830-845	QYICAKNSKEGITE
	DR12	DRB1*1201	Herauf	unknown eluted peptide	EAIIQLKINPYLS
	DR13	DRB1*1302	H0301	Tet. tox. 830-845 S>A	QYICAKNAKFGITE
	DR51	DRB5*0101	GM3107 or L416.3	HA 307-319	QYICAKNAKFGITE
	DR52	DRB5*0201	MAT	HA 307-319	PXYKQNTLKLAT
	DR53	DRB5*0101	L255.1	Tet. tox. 1272-1284	NGQIGNPDRIL
	DR53.1	DRB5*0101	L257.6	non-natural (717.01)	YARFQSTTLKQKT
	DR53.1	DRB5*0101	PF	non-natural (ROIV)	YAHAAHAAHAAHAA
Mouse	IA ^b	D ² A1*0301/DQ81*0301	DB27.4	non-natural (ROIV)	YAHAAHAAHAAHAA
	IA ^d		A20	non-natural (ROIV)	YAHAAHAAHAAHAA
	IA ^e		CH-12	HEL 46-61	YNTDGSYDYLQNSR
	IA ^f		LS102.9	non-natural (ROIV)	YAHAAHAAHAAHAA
	IA ^g		91.7	non-natural (ROIV)	YAHAAHAAHAAHAA
	IE ^d		A20	Lambda repressor 12-26	YLEDAREKKAIYEKKK
	IE ^k		CH-12	Lambda repressor 12-26	YLEDAREKKAIYEKKK

Table XXV. Monoclonal antibodies used in MHC purification.

Monoclonal antibody	Specificity
W6/32	HLA-class I
B123.2	HLA-B and C
IVD12	HLA-DQ
LB3.1	HLA-DR
M1/42	H-2 class I
28-14-8S	H-2 D ^b and L ^d
34-5-8S	H-2 D ^d
B8-24-3	H-2 K ^b
SF1-1.1.1	H-2 K ^d
Y-3	H-2 K ^b
10.3.6	H-2 I A ^k
14.4.4	H-2 IE ^d , IE ^k
MKD6	H-2 I A ^d
Y3JP	H-2 I A ^b , I A ^s , I A ^u

Table XXVI: HCV-derived conserved high algorithm A*0201-binding peptides

Peptide	Molecule	1st Position	Sequence	Consv.	A2-super-type binding capacity (IC50 nM)					
					A*0201	A*0202	A*0203	A*0206	A*6802	A2 XRN
1073.05	NS4	1812	LLFNILGWV	85	4.2	113	3.2	19	33	5
1090.18	NS1/E2	728	FLLADARV	92	18	90	149	247	111	5
1013.02	NS4	1590	YLVAYQATV	85	20	39	16	82	33	5
1090.22	NS5	2611	RLVFPDLGV	79	56	391	10	370	8000	4
1013.1002	CORE	132	DLMGYIPLV	79	80	4778	204	481	12	4
24.0073	NS4	1920	WMNRLIAFA	100	122	130	3.3	1609	400	4
24.0075	NS4	1666	VLVGGVLA	85	185	331	32	308	3077	4
1174.08	NS4	1769	HMWNFISGI	92	15	10750	77	132	7547	3
1073.06	NS4	1851	ILAGYGAGV	79	116	143	5.0	755	889	3
1073.07	CORE	35	YLLPRGPRRL	92	125	6143	455	416	10256	3
24.0071	NS1/E2	726	LLFLLADA	100	217	287	455	3364	3077	3
1.0119	LORF	1131	YLVTRHADV	85	455	2048	3.6	71	3077	3
24.0065	NS4	1891	ILSPGALVV	92	238	10750	27	1028	3077	2
1013.12	NS1/E2	686	ALSTGLIHL	85	313	7167	45	18500	10256	2
939.14	NS1/E2	696	HLHQNIYDV	85	500	3071	19	1370	10811	2
1090.21	NS5	2918	RLHGLSAFSL	79	179	782	625	18500	12500	1

Table XXVII: HCV-derived conserved high algorithm A*03 and/or A*11 binding peptides

Peptide	Molecule	1st Position	Sequence	Consv.	A3-superotype binding capacity (IC50 nM)										A3 XRN
					A*03	A*11	A*3101	A*3301	A*6801	A3					
1.0952	CORE	51	KTSESRQR	92	69	94	67	1813	145	4					
1073.11	CORE	43	RLGVRATRK	79	12	207	429	-	-	3					
1.0955	ENV1	290	QLFTSPRR	79	15	182	621	3766	3	3					
1073.13	NS1/E2	632	RMVVGVEHR	100	15	300	95	9667	1778	3					
1.0123	NS3	1396	LIFCHSKKK	100	20	32	2535	24167	333	3					
1073.10	NS4	1863	GVAGALVAFK	85	28	4	3273	26364	118	3					
24.0090	NS4	1864	VAGALVAFK	85	46	7	3750	11600	258	3					
24.0086	NS3	1262	LGFGAYMSK	85	136	21	2950	22308	222	3					
1174.16	NS1/E2	557	WMNSTGFTK	79	208	74	12857	690	1429	2					
1073.14	NS3	1261	TLFGAYMSK	85	136	98	-	22308	8889	2					
1090.23	LORF	1183	AVCTRGVAK	79	423	240	16364	-	-	2					
1090.24	NS5	2596	EVFCVQPEK	85	13750	222	-	-	18	2					
24.0103	NS1/E2	647	AACNWTGRER	85	36667	429	400	5273	4444	2					
1073.16	NS3	1232	HLHAPTGSQK	85	19	2500	-	-	2857	1					
1073.12	NS3	1395	HLFCHSKKK	100	423	-	20000	-	-	1					
1090.26	NS3	1395	HLFCHSKKK	100	440	10000	-	-	8000	1					

* A dash indicates IC50nM > 30,000

Table XXVIII: HCV derived conserved B*0702 binding peptides

A. High conservancy 9- and 10-mer peptides.

Peptide	Molecule	1st Position	Sequence	Consv.	B7-supertype binding capacity (IC50 nM)						
					B*0702	B*5301	B*51	B*5301	B*5401	B7 XRN	
1145.12	Core	169	LPGCSFSIF	92	28	90	100	114	6667	4	
15.0048	E2	681	LPALSTGLI	85	157	-	2.8	1500	20000	2	
15.0234	NS3	1620	KPTLHGPTPL	79	3.9	-	27500	-	-	1	
15.0247	NS5	2835	APTLWARMIL	79	6.3	-	5500	-	-	1	
15.0042	CORE	99	SPRGSRPSW	79	14	-	11000	-	-	1	
15.0039	Core	57	QPRGRQPI	92	24	-	-	-	-	1	
15.0218	Core	37	LPKRGRLGV	92	29	-	6111	-	4000	1	
15.0060	NS5	2615	SPGQVFEFL	79	46	-	27500	-	-	1	
15.0043	Core	111	DPREBSRNL	85	324	-	-	-	-	1	
15.0063	NS5	2835	APTLWARMIL	79	344	-	4583	-	-	1	
1292.17	NS5	2317	PPVWHGCP	79	393	-	-	-	-	1	
15.0239	NS4	1893	SPGALVGVV	79	423	-	3438	-	-	1	
15.0235	NS3	1621	TPLLRYLGAV	92	458	-	6875	-	909	1	

Table XXVIII: HCV derived conserved B*0702 binding peptides

B. Additional HCV derived B7 supermotif peptides.				B7-superotype binding capacity (IC50 nM)					
Peptide	Molecule	1st Position	Sequence	Consv.	B*0702	B*1501	B*51	B*5301	B*5401
29.0035	NS3	1378	IPFYGKAI	92	458	-	46	-	50
29.0040	Core	37	LPRRGRL	92	0.85	-	306	-	5000
29.0036	Core	137	IPLVGAFL	79	13	2250	79	-	2857
16.0187	NS1/E2	680	LPGCSFTLPA	64	423	24000	9167	-	15
29.0039	Core	169	LPGCSFSI	92	500	200	932	620	6250
15.0219	Core	142	APLGAARAL	71	9.5	-	-	-	12500
29.0031	NS5	2869	APTLWARM	79	13	-	4583	-	4348
15.0231	NS3	1512	RPSGMFDSV	71	153	-	-	-	-
29.0085	NS5	2474	LPINALNSL	57	220	18000	1170	-	11111
29.0037	NS5	2608	KPARLIVF	85	367	-	3235	-	16667
15.0237	NS4	1789	NPALASLMF	71	393	9000	5000	-	-
29.0118	NS5	2869	APTLWARMILM	79	423	-	-	-	3030
29.0042	NS4	1720	LPYIEQGM	85	423	-	1375	-	7692

C. Engineered analogs of B7 supermotif peptides.

C. Engineered analogs of B7 supermotif peptides.				B7-superotype binding capacity (IC50 nM)					
Peptide	Molecule	1st Position	Sequence	Consv.	B*0702	B*1501	B*51	B*5301	B*5401
1145.12	Core	169	LPGCSFSIF	92	28	90	100	114	6667
1292.24	Core	169	LPGCSFSII	37	4364	5.3	262	1056	3
1145.13	Core	169	FPGCSFSIF	19	1.6	132	3.2	6.7	5

* A dash indicates IC50 nM >30,000.

Table XXIX: HCV-derived A1- and A24-motif containing peptides

A. A1-motif peptides

Peptide	Molecule	Position	Sequence	Conserv.	HLA-A*0101 binding (IC50 nM)
13.0019	NS5	2922	LSAFSLHSY	79	31
1.0509	NS5	2921	GLSAFSLHSY	79	61
1069.62	NS3	1128	CTCGSSDLY	79	68
24.0093	NS5	2129	EVDGVRHLRY	100	167
13.0016	NS3	1241	KSTKVPAAAY	85	1923
1.0125	NS3	1525	CYDAGCAWY	79	4032
24.0008	E1	206	DCSNSSIVY	85	16667
24.0094	NS5	2720	TNSKGQNCGY	100	-
24.0096	NS3	1240	GKSTKVPAAAY	85	-
24.0100	NS3	1292	TGAPITYSTY	85	-
	NS3	1263	VAATLGFAGY	100	-
	NS5	2639	VMGSSYGFQY	79	-
	NS5	2640	MGSSYGFQY	79	-

A dash indicates IC50 nM >25000

B. A24 -motif peptides

Peptide	Molecule	Position	Sequence	Conserv.	HLA-A*2402 binding (IC50 nM)
24.0092	NS4	1763	FWAKHMMWNF	85	1.7
13.0075	NS4	1778	QYLAGLSTL	100	250
1073.18	NS1/E2	636	MYVGGVEHRL	92	444
13.0074	NS3	1297	TYSTYGFKL	85	522
13.0134	NS5	2647	QYSPGQRVEF	79	667
24.0091	NS4	1772	NFISGIQYL	100	706
13.0131	Core	135	GYIPLVGAPL	79	2105
24.0108	Core	173	SFSIFLLALL	100	2927
13.0132	NS3	1248	AYAAQGYKVL	79	13333
13.0133	NS4	1859	GYGAGVAGAL	85	-
1174.08	NS4	1769	HMWNFISGI	93	-
	E1	317	RMAWDMMMNW	85	-
	NS1/E2	635	RMVYGGVEHRL	93	-
	NS3	1422	YYRGLDVSVI	100	-
	NS3	1468	DFSLDPTFTI	100	-
	NS3	1608	SWDQMWKCL	79	-
	NS3	1664	TWVVLGGVL	85	-
	NS4	1732	QFKQKALGL	85	-
	NS4	1732	QFKQKALGLL	85	-
	NS4	1765	FWAKHMMWNFI	85	-
	NS4	1919	QWMNRLIAP	100	-
	NS5	2241	LWRQEMGGNI	85	-
	NS5	2669	GFSDYDTRCF	79	-
	NS5	2875	RMILMTHTFF	85	-

A dash indicates IC50 nM >25000

Table XXX: Immunogenicity of A2-supertype cross-reactive binders

Peptide	Sequence	Protein	Position	Immunogenicity				Frequency	Response	
				Barnaba; Barnaba; patients contacts	Human ^a	Transgenic mice ^b				
1073.05	LLFNILGGVV	NS4	1812	1/6	7/17	2/21	0/6	10/50	6/6	6.4 (1.7)
1090.18	FLLLADARV	NS1/E2	728	2/6	7/17	1/21	0/6	10/50	5/6	9.5 (3.0)
1013.02	YLVAYQATV	NS4	1590	1/6	4/17	1/21	0/6	6/50	5/6	8.5 (3.7)
1090.22	RLVFPDLGV	NS5	2578	2/6	5/17	0/21	0/6	7/50	0/6	-
1013.1002	DLMGYIPLV	Core	132	2/6	7/17	1/21	1/6	11/50	5/6	8.8 (2.6)
24.0073	WMNRLIAFA	NS4	1920	1/6	3/17	2/21	1/6	7/50	0/6	-
24.0075	VLVGGVLA	NS4	1666	1/6	6/17	3/21	1/6	11/50	0/6	-
1174.08	HMWNFSIGI	NS4	1769	3/6	3/17	2/21	0/6	8/50	6/6	6.4 (1.7)
1073.06	ILAGYGAGV	NS4	1851	2/6	3/17	0/21	0/6	5/50	3/6	54.7 (3.3)
1073.07	YLLPRGPRLL	CORE	35	2/6	5/17	7/21	1/6	17/50	4/6	59.1 (7.2)
24.0071	LLFLLADA	NS1/E2	726	2/6	9/17	0/21	0/6	11/50	0/6	-
1.0119	YLVTRHADV	NS3	1131	6/6	10/17	0/21	1/6	17/50	0/6	-

a. Data shown represents the number of positive responses over the total number of patients or contacts examined.

b. Frequency represents the number of positive responses over the total number of mice examined. Response indicates the average magnitude (standard deviation) of the response in positive animals, measured in lytic units.

Table XXXI: Immunogenicity of A3-supertype cross-reactive binders

Peptide	Sequence	Protein	Position	Immunogenicity					Frequency	Response
				Human ^a				overall		
				Barnaba patients	Barnaba contacts	Chisari	Pape			
1.0952	KTSSRSQPR	CORE	51	2/16	1/4	3/12	0/6	6/38	3/6	23.4 (1.3)
1073.11	RLGVATRRK	CORE	43	4/16	1/4	7/12	1/6	13/38	3/6	42.2 (1.2)
1.0955	QLFTFSRR	ENV	290	1/16	0/4	6/12	1/6	8/38		
1073.13	RMVVGVEHR	NS1/E2	632	5/16	1/4	4/12	1/6	11/38	2/6	2.8 (1.1)
1.0123	LIFCHSKKK	NS3	1396	6/16	1/4	4/12	2/6	13/38	3/6	4.4 (1.1)
1073.10	GVAGALVAFK	NS4	1863	3/16	0/4	6/12	2/6	11/38	6/6	56.5 (1.7)
24.0090	VAGALVAFK	NS4	1864	4/16	1/4	6/12	0/4	11/38	1/6	7.1
24.0086	TLGFGAYMSK	NS3	1262	6/16		2/12	2/5	10/33		

a. Data shown represents the number of positive responses over the total number of patients or contacts examined.

b. Frequency represents the number of positive responses over the total number of mice examined. Response indicates the average magnitude (standard deviation) of the response in positive animals, measured in lytic units.

Table XXXII. Candidate HCV-derived HTL epitopes

Selection criteria	Peptide	Sequence	Source	Conservancy	
				Total	Core
A. DR-supermotif conserved 15mers	1283.01	GQVGGVYLLPRRGPR	HCV Core 28	93	93
	1283.02	VYLLPRRGPRGLGVRA	HCV Core 34	93	93
	1283.03	GWLLSPRGSRPSWGPT	HCV Core 95	79	79
	1283.04	LKGVIDTLTCGFADL	HCV Core 119	79	86
	1283.05	IDTLTCGFADLMGYI	HCV Core 123	86	86
	1283.06	ADLMGYIPLVGAPLG	HCV Core 131	79	79
	1283.07	GVRVLEDGVNYYATGN	HCV Core 154	86	86
	1283.08	GVNYYATGNLPGCSFS	HCV Core 161	79	86
	1283.09	GCSSIFLLALLSCL	HCV Core 171	86	100
	1283.10	GHRMAWDMMNMWSPT	HCV E1 315	86	86
	1283.11	CGPVYCTSPVVVG	HCV NS1/E2 306	93	93
	1283.12	VYCFTSPVVVGTTD	HCV NS1/E2 309	93	93
	1283.13	GNWFGCTWMNSTGET	HCV NS1/E2 350	79	86
	1283.14	FTTLPALSTGLIHL	HCV NS1/E2 684	79	86
	1283.17	DLYLVTRHADVIPVR	HCV NS3 1134	79	79
	1283.18	RAAVCTRGVAKAVDF	HCV NS3 1186	79	79
	1283.20	AQGYKVLVLNPSVAA	HCV NS3 1251	79	100
	1283.21	GYKVLVLNPSVAATL	HCV NS3 1253	100	100
	1283.22	VLVLNPSVAATLGF	HCV NS3 1256	100	100
	1283.23	GTVLDAQETAGARLV	HCV NS3 1335	86	86
	1283.24	GARLVVLATATPPGS	HCV NS3 1345	79	86
	1283.25	GRHLFCHSKKKCDE	HCV NS3 1393	100	100
	1283.27	DSVIDCNTCVTQTV	HCV NS3 1454	86	86
	1283.28	TVDFSLDPTFTIET	HCV NS3 1466	79	100
	1283.30	FTGLTHIDAHFLSQ	HCV NS3 1567	93	93
	1283.31	YLVAYQATVCARAQA	HCV NS3 1591	79	93
	1283.32	KPTLHGPTLLYRLG	HCV NS4 1620	79	79
	1283.33	LEVVTSTWVLVGGVL	HCV NS4 1658	86	86
	1283.34	TWVLVGGVLAALAA	HCV NS4 1664	86	86
	1283.35	AEQFKQKALGLLQTA	HCV NS4 1730	86	86
	1283.40	PAISPGALVVGVCVA	HCV NS4 1889	79	93
	1283.41	GALVVGVCVAAILRR	HCV NS4 1895	79	79
	1283.42	CAAILRRHVGPGEA	HCV NS4 1903	79	79
	1283.43	AVQWMNRLIAFASRG	HCV NS4 1917	100	100
	1283.44	MNRLIAFASRGHVS	HCV NS4 1921	86	100
	1283.48	ANLLWRQEMGGNTR	HCV NS5 2238	86	86
	1283.49	RQEMGGNTRVSEEN	HCV NS5 2243	86	86
	1283.52	ARLVFPDLGVRVCE	HCV NS5 2610	79	79
	1283.53	FPDLGVRVCEKMALY	HCV NS5 2615	79	100
	1283.54	GVRVCEKMALYDVVS	HCV NS5 2619	79	100
	1283.56	QPEYDLELITSCSSN	HCV NS5 2808	79	93
	1283.57	LELITSCSSNVSAH	HCV NS5 2813	79	100
	1283.58	PTLWARMILMTHFFS	HCV NS5 2870	79	86
	1283.59	LHGLSAFSLHSYSPG	HCV NS5 2919	79	79
	1283.60	AFSLHSYSPGEINRV	HCV NS5 2924	79	79

Table XXXII. Candidate HCV-derived HTL epitopes

Selection criteria	Peptide	Sequence	Source	Conservancy	
				Total	Core
B. High algorithm conserved core	1283.15	VVLLFLLADARVCS	HCV NS1/E2 724	29	100
	1283.16	SKGWRLLAPITAYAQ	HCV NS3 1025	29	79
	1283.19	PQTFQVAHLHPTGS	HCV NS3 1225	43	85
	1283.26	DVVVVATDALMTGYT	HCV NS3 1436	43	79
	1283.29	WESVFTGLTHIDAHF	HCV NS3 1563	43	92
	1283.45	LTSMLTDPSSHIAET	HCV NS5 2176	57	100
	1283.46	ASQLSAPSLKATCTT	HCV NS5 2208	50	79
	1283.47	DADLJEANLLWRQEM	HCV NS5 2232	50	85
	1283.50	SYTWGTALITPCAAE	HCV NS5 2456	64	79
	1283.51	TTIMAKNEVFCVQPE	HCV NS5 2589	64	85
	1283.55	OSSYGFQYSPGQKVE	HCV NS5 2641	71	79
	1283.61	ASCLRKLGVPLRLVW	HCV NS5 2939	50	85
C. Collaborator	P098.03	AAYAAQGYKVLVLNPSVAAT	HCV NS3 1242-1261	71	100
	P098.04	GYKVLVLNPSVAATLFGGAY	HCV NS3 1248-1267	100	
	P098.05	GYKVLVLNPSVAAT	HCV NS3 1248-1261	100	
	F134.01	RRPDQVKFPGGGQVGGVY	HCV Core 17-35	86	
	F134.02	DVKFPGGGQIVGGVYLLPRR	HCV Core 21-40	86	
	F134.03	GYKVLVLNPSVAATLFGGAY	HCV NS3 1253-1272	100	
	F134.04	TLHGPTLLYRLGAVQNEIT	HCV NS4 1622-1641		79
	F134.05	NFISGIQYLGLSTLPGNPA	HCV NS4 1772-1791	100	
	F134.06	LLFNILGGWVAAQLAAPGAA	HCV NS4 1812-1831		86
	F134.07	PGPEGAVQWMNRLIAFASRG	HCV NS4 1912-1931	86	100
	F134.08	GEGAVQWMNRLIAFASRGNHV	HCV NS4 1914-1934	100	
	Page 21	AIPLFVKGGRHLIFCHSKR	HCV NS3 1379-1398	21	100
D. DR3 motif	Page 22	GRHLIFCHSKRKDELATKL	HCV NS3 1388-1407		100
	Page 29	SVIDCNTCVTQTVDFSLDPT	HCV NS3 1450-1469	86	
	35.0102	GVRVLEDGVNYYATGN	HCV 154	86	86
	35.0103	SAMYYGDLGGSVFLV	HCV 273	57	86
	35.0104	GHEKAYWDAAQNNVSP	HCV 315	86	86
	35.0105	SDLVLYTRHADVIPV	HCV 1133	79	86
	35.0106	VVVVATDALMTGYTG	HCV 1437	42	86
	35.0107	TVDFSLDPTFTIETT	HCV 1466	79	100
	35.0108	DSSVLCECYDAGCAW	HCV 1518	71	93
	35.0109	GLPVCQDHLFEWESV	HCV 1552	42	86
	35.0110	GMQLAEQFKQKALGL	HCV 1726	57	86
	35.0111	PTHYVPESDAARVT	HCV 1936	86	86
	35.0112	GSQLPCEPEDVAVL	HCV 2162	64	86
	35.0113	LTSMLTDPSSHIAET	HCV 2176	57	100
	35.0114	MPPLEGPGDPLSD	HCV 2401	79	100
	35.0115	QPEYDLELITSCSN	HCV 2808	79	93
	1283.25	GRHLIFCHSKKKCDE	HCV NS3 1393-1407		

Table XXXIII. HLA-DR screening panels

Screening Panel	Antigen	Representative Assay		Phenotypic Frequencies						
		Alleles	Allele	Alias	Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.
Primary	DR1	DRB1*0101-03	DRB1*0101	(DR1)	18.5	8.4	10.7	4.5	10.1	10.4
	DR4	DRB1*0401-12	DRB1*0401	(DR4w4)	23.6	6.1	40.4	21.9	29.8	24.4
	DR7	DRB1*0701-02	DRB1*0701	(DR7)	26.2	11.1	1.0	15.0	16.6	14.0
	Panel total				59.6	24.5	49.3	38.7	51.1	44.6
Secondary	DR2	DRB1*1501-03	DRB1*1501	(DR2w2.81)	19.9	14.8	30.9	22.0	15.0	20.5
	DR2	DRB5*0101	DRB5*0101	(DR2w2.82)	-	-	-	-	-	-
	DR9	DRB1*0901.09012	DRB1*0901	(DR9)	3.6	4.7	24.5	19.9	6.7	11.9
	DR13	DRB1*1301-06	DRB1*1302	(DR6w19)	21.7	16.5	14.6	12.2	10.5	15.1
	Panel total				42.0	33.9	61.0	48.9	30.5	43.2
Tertiary	DR4	DRB1*0405	DRB1*0405	(DR4w15)	-	-	-	-	-	-
	DR8	DRB1*0801-5	DRB1*0802	(DR8w2)	5.5	10.9	25.0	10.7	23.3	15.1
	DR11	DRB1*1101-05	DRB1*1101	(DR3w11)	17.0	18.0	4.9	19.4	18.1	15.5
	Panel total				22.0	27.8	29.2	29.0	39.0	29.4
Quaternary	DR3	DRB1*0301-2	DRB1*0301	(DR3w17)	17.7	19.5	0.4	7.3	14.4	11.9
	DR12	DRB1*1201-02	DRB1*1201	(DR5w12)	2.8	5.5	13.1	17.6	5.7	8.9
	Panel total				20.2	24.4	13.5	24.2	19.7	20.4

Table XXXIV. HLA-DR binding capacity of target derived peptides: DR-supermotif and algorithm positive peptides.

Peptide	Sequence	Source	Binding capacity (IC50 nM)										DR alleles bound
			DR1	DR2w281	DR2w282	DR4w4	DR4w15	DR5w11	DR6w19	DR7	DR8w2	DR9	Lab
1283.21	AAVAAQGYKVLVLPNSVAATLQFGAY	HCV NS3 1242-1267	4.5	350		5.2	567	143	5.1	89	288	54	175
1283.20	GYKVLVLPNSVAATL	HCV NS3 1253				7.9	224	74	3.9	833	175	375	298
F98.03	AQGYKVLVLPNSVAA	HCV NS3 1251	6.0	650									
F98.05	AAVAAQGYKVLVLPNSVAAT	HCV NS3 1242	2.9	48	483	18	1234	103	11	96	60	240	9
F98.03	GYKVLVLPNSVAAT	HCV NS3 1248-1261	1.4	39	5695	7.8	141	75	3.5	126	21	266	9
F98.04	GYKVLVLPNSVAATLQFGAY	HCV NS3 1248-1267	3.5	42	8154	9.7	1500	240	4.1	23	80	20	8
	GEGAVQWNNRLIAFASRGNHVS	HCV NS4 1914-1935											
1283.44	MNNRLIAFASRGNHVS	HCV NS4 1921	66	4.8				585	45	7.3	227	102	313
F134.08	GEGAVQWNNRLIAFASRGNHV	HCV NS4 1914	3.2	182	361			345	221	158	4818	147	8
1283.16	SKGWRLIAATAYAQ	HCV NS3 1025	0.36	125	23	24	152	4.8	962	54	384	8	
1283.55	GSSYGFQYSPGQKVE	HCV NS5 2641	11	667	417	745	20000	19	156	68	571	7	
1283.61	ASCLRLKGLVPLRVW	HCV NS5 2939	5.0	16	217	6250	78	645	862	671	862	7	
F134.05	NFISGQYLAGLSTLPNGPA	HCV NS4 1772	10	606	84			29		70	441	6	

Shading indicates IC50 > 1 μM.

A dash (-) indicates IC50 > 20 μM.

Table XXXV. HLA-DR binding capacity of 3 DR3 motif-containing peptides

Peptide	Sequence	Source	DR3 binding (IC50 nM)
35.0106	VVVVATDALMTGYTG	HCV 1437	427
35.0107	TVDFSLDPTFTIET	HCV 1466	235
1283.25	GRHLFCHSKKKCDE	HCV NS3 1393	ND

Table XXXVIa: HCV-derived CTL epitope candidates

Peptide	Molecule	1st Position	Sequence	Consv.	Selection criteria
1073.05	NS4	1812	LLFNILGGWV	85	A2-supertype
1090.18	NS1/E2	728	FLLLADARV	92	A2-supertype
1013.02	NS4	1590	YLVAQATV	85	A2-supertype
1090.22	NS5	2611	RLIVPDLGV	79	A2-supertype
1013.1002	CORE	132	DLMGYIPLV	79	A2-supertype
24.0073	NS4	1920	WMNRLIAFA	100	A2-supertype
24.0075	NS4	1666	VLVGGVLA	85	A2-supertype
1174.08	NS4	1769	HMWNFISGI	92	A2-supertype
1073.06	NS4	1851	ILAGYGAGV	79	A2-supertype
1073.07	CORE	35	YLLPRGPRLL	92	A2-supertype
24.0071	NS1/E2	726	LLFLLADA	100	A2-supertype
1.0119	LORF	1131	YLVTRHADV	85	A2-supertype
1.0952	CORE	51	KTSEKSPQR	92	A3-supertype
1073.11	CORE	43	RLGVRAIRK	79	A3-supertype
1.0955	ENV1	290	QLFTFSPPR	79	A3-supertype
1073.13	NS1/E2	632	RMVVGVEHR	100	A3-supertype
1.0123	NS3	1396	LIFCHSKKK	100	A3-supertype
1073.10	NS4	1863	GVAGALVAFK	85	A3-supertype
24.0090	NS4	1864	VAGALVAFK	85	A3-supertype
24.0086	NS3	1262	TLGFQAYMSK	85	A3-supertype
F104.01	NS5	3003	VGHYLLPNR	79	A31
1145.12	Core	169	LPGCSFSIF	92	B7-supertype
29.0035	NS3	1378	IPFYGKAI	92	B7-supertype
13.0019	NS5	2922	LSAFSLHSY	79	A1
1069.62	NS3	1128	CTCGSSDLY	79	A1
24.0092	NS4	1765	FWAKHWNF	85	A24

Table XXXVlb: HCV-derived HTL epitope candidates

Region	Peptide	Motif ¹	Sequence
HCV NS3 1025-1039	1283.16	DR	SKGWRLAPITAYAQ\
HCV NS3 1242-1267	F98.03	DR	AAYAAQGYKVLNFSVAAT.
HCV NS3 1393-1407	1283.25	DR3	GRHLFCHSKKKCDE.
HCV NS3 1437-1451	35.0106	DR3	VVVVAITDALMTGYTG.
HCV NS3 1466-1480	35.0107	DR3	TVDFSLDPTFTIETT.
HCV NS4 1772-1790	F134.05	DR	NFISGIQYLAGLSTLPGNPA\
HCV NS4 1914-1935	F134.08	DR	GECAVQWMNELLAFASRGNHV\
HCV NS5 2641-2655	1283.55	DR	GSSYGFQYSPGQRVE.
HCV NS5 2939-2953	1283.61	DR	ASCLRLKLGVPPLRVW\

1. Peptides identified on the basis of either the DR PL-P6 supermotif or by use of the DR1-4-7 algorithms are indicated by 'DR'. Peptides identified using the DR3 motif are indicated by 'DR3'.

Table XXXVII. Estimated population coverage by a panel of HCV derived HTL epitopes

Antigen	Alleles	Representative assay	No. of epitopes ²	Population coverage (phenotypic frequency)						
				Cauc.	Blk.	Jpn.	Chn.	Hisp.	Avg.	
DR1	DRB1*0101-03	DR1	6	18.5	8.4	10.7	4.5	10.1	10.4	
DR2	DRB1*1501-03	DR2w2 B1	3	19.9	14.8	30.9	22.0	15.0	20.5	
DR2	DRB5*0101	DR2w2 B2	6	-	-	-	-	-	-	
DR3	DRB1*0301-2	DR3	2	17.7	19.5	0.40	7.3	14.4	11.9	
DR4	DRB1*0401-12	DR4w4	5	23.6	6.1	40.4	21.9	29.8	24.4	
DR4	DRB1*0401-12	DR4w15	3	-	-	-	-	-	-	
DR7	DRB1*0701-02	DR7	5	26.2	11.1	1.0	15.0	16.6	14.0	
DR8	DRB1*0801-5	DR8w2	5	5.5	10.9	25.0	10.7	23.3	15.1	
DR9	DRB1*09011,09012	DR9	3	3.6	4.7	24.5	19.9	6.7	11.9	
DR11	DRB1*1101-05	DR5w11	5	17.0	18.0	4.9	19.4	18.1	15.5	
DR13	DRB1*1301-06	DR6w19	2	21.7	16.5	14.6	12.2	10.5	15.1	
Total				98.5	95.1	97.1	91.3	94.3	95.1	

1. Total population coverage has been adjusted to account for the presence of DRX in many ethnic populations. It has been assumed that the range of specificities represented by DRX alleles will mirror those of previously characterized HLA-DR alleles. The proportion of DRX incorporated under each motif is representative of the frequency of the motif in the remainder of the population. Total coverage has not been adjusted to account for unknown gene types.

2. Number of epitopes represents a minimal estimate, considering only the epitopes shown in Table 6. Additional alleles possibly bound by nested epitopes have not been accounted.

TABLE Ia

SUPERMOTIFS	POSITION 2 (Primary Anchor)	POSITION 3 (Primary Anchor)	POSITION C Terminus (Primary Anchor)
A1	T, I, L, V, M, S		F, W, Y
A2	V, Q, A, T		I, V, L, M, A, T
A3	V, S, M, A, T, L, I		R, K
A24	Y, F, W, I, V, L, M, T		F, I, Y, W, L, M
B7	P		V, I, L, F, M, W, Y, A
B27	R, H, K		F, Y, L, W, M, I, V, A
B58	A, T, S		F, W, Y, L, I, V, M, A
B62	Q, L, I, V, M, P		F, W, Y, M, I, V, L, A
MOTIFS			
A1	T, S, M		Y
A1		D, E, A, S	Y
A2.1	V, Q, A, T*		V, L, I, M, A, T
A3.2	L, M, V, I, S, A, T, F, C, G, D		K, Y, R, H, F, A
A11	V, T, M, L, I, S, A, G, N, C, D, F		K, R, H, Y
A24	Y, F, W		F, L, I, W

*If 2 is V, or Q, the C-term is not L

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

WHAT IS CLAIMED IS

1. A composition comprising a prepared hepatitis C virus (HCV) epitope consisting of an amino acid sequence selected from the group consisting of:

FLLADARV,	YLVAYQATV,	RLIVFPDLGV,
DLMGYIPLV,	WMNRLIAFA,	VLVGGVLA,
HMWNFISGI,	ILAGYGAGV,	YLLPRRGPR,
LLFLLADA,	YLVTRHADV,	KTTERSQPR,
RLGVRATRK,	QLFTFSPRR,	RMVVGVEHR,
LIFCHSKKK,	GVAGALVAFK,	VAGALVAFK,
TLGFGAYMSK,	LPGCSFSIF,	LSAFSLHSY,
CTCGSSDLY,	FWAKHMWNF,	SKGWRLAPITAYAQ,
AAAYAAQGYKVLNPSVAAT,	GRHLIFCHSKKKCDE,	VVVVATDALMTGYTG,
TVDFSLDPTFTIETT,	NFISGIQYLAGLSTLPGNPA,	
GEGAVQWMNRLIAFASRGNHV,	GSSYGFQYSPGQQRVE,	ASCLRKLGPPLRVW,
and LTCGFADLMGY.		

2. The composition of claim 1, further comprising two epitopes selected from the group in claim 1.

3. The composition of claim 2, further comprising three epitopes selected from the group in claim 1.

4. The composition of claim 1, wherein the composition further comprises a CTL epitope selected from the group consisting of LTDPSHITA, LADGGCSGGAY, RMILMTHFF, VMGSSYGF, FWAKHMWNFI, LLFNILGGWV, IPFYGKAI, and VGIYLLPNR.

5. The composition of claim 1, wherein the composition further comprises an HTL epitope.

6. The composition of claim 5, wherein the HTL epitope is a pan DR binding molecule.

7. The composition of claim 1, wherein the epitope is on or within a liposome.
8. The composition of claim 1, wherein the peptide is joined to a lipid.
9. The composition of claim 1, wherein the epitope is bound to an HLA heavy chain, β 2-microglobulin, and strepavidin complex, whereby a tetramer is formed.
10. The composition of claim 1, wherein the epitope is bound to an HLA molecule on an antigen presenting cell.
11. The composition of claim 10, wherein the antigen presenting cell is a dendritic cell.
12. The composition of claim 1, the composition further comprising a pharmaceutical excipient.
13. The composition of claim 1, further wherein the epitope is in a unit dose form.
14. A composition comprising a prepared peptide of less than 250 amino acid residues comprising at least two hepatitis C virus (HCV) peptide epitopes selected from the group consisting of:
- | | | |
|-------------|-------------|-----------------|
| FLLADARV, | YLVAYQATV, | RLVFPDLGV, |
| DLMGYIPLV, | WMNRLIAFA, | VLVGGVLAA, |
| HMWNFISGI, | ILAGYGAGV, | YLLPRRGPR, |
| LLFLLADA, | YLVTRHADV, | KTSERSQPR, |
| RLGVRATRK, | QLFTFSPRR, | RMVYGGVEHR, |
| LIFCHSKKK, | GVAGALVAFK, | VAGALVAFK, |
| TLGFGAYMSK, | LPGCSFSIF, | LSAFSLHSY, |
| CTCGSSDLY, | FWAKHMWNF, | SKGWRLAPITAYAQ, |

AAAYAAQGYKVLVLPNSVAAT, GRHLIFCHSKKKCDE, VVVVATDALMTGYTG,
TVDFSLDPTFTIETT, NFISGIQYLAGLSTLPGNPA,
GEGAVQWMNRLIAFASRGNHV, GSSYGFQYSPGQRVE, ASCLRKLGPVPLRVW,
and LTCGFADLMGY.

15. The composition of claim 14, wherein at least two epitopes are linked via a spacer.
16. The composition of claim 14, further comprising a third epitope.
17. The composition of claim 16, wherein the third epitope is selected from the group consisting of LTDPSHITA, LADGGCSGGAY, RMILMTHFF, VMGSSYGF, FWAKHMWNFI, LLFNILGGWV, IPFYGKAI, and VGIYLLPNR.
18. The composition of claim 16, further comprising a third epitope that is an HTL epitope.
19. The composition of claim 18, wherein the HTL epitope is a panDR binding molecule.
20. The composition of claim 14, wherein the peptide is on or within a liposome.
21. The composition of claim 14, wherein the peptide is joined to a lipid.
22. The composition of claim 14, wherein the peptide further comprises at least three of the epitopes in the group of claim 14.
23. The composition of claim 14, wherein the peptide further comprises at least four of the epitopes in the group of claim 14.
24. The composition of claim 14, wherein the peptide further comprises at least five of the epitopes in the group of claim 14.

25. The composition of claim 14, wherein the peptide further comprises at least six of the epitopes in the group of claim 14.

26. The composition of claim 14, the composition further comprising a pharmaceutical excipient.

27. The composition of claim 14, further wherein the epitope is in a unit dose form.

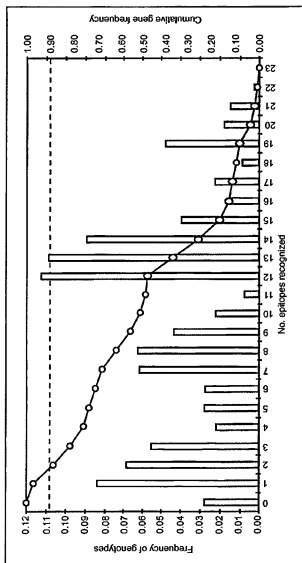
28. A composition comprising at least six prepared HCV epitopes each consisting of an amino acid sequence selected from the group consisting of:

FLLLADARV,	YLVAYQATV,	RLIVFPDLGV,
DLMGYIPLV,	WMNRLIAFA,	VLVGVLAA,
HMWNFISGI,	ILAGYGAGV,	YLLPRRGPR,
LLFLLADA,	YLVTRHADV,	KTTERSQPR,
RLGVRATRK,	QLFTFSPRR,	RMVVGVEHR,
LIFCHSKKK,	GVAGALVAFK,	VAGALVAFK,
TLGFGAYMSK,	LPGCSFSIF,	LSAFSLHSY,
CTCGSSDLY,	FWAKHMWNF,	SKGWRLLAPITAYAQ,
AAAYAAQGYKVLVNLPSVAAT,	GRHLIFCHSKKKCDE,	VVVVATDALMTGYTG,
TVDFSLDPTFTIETT,	NFISGIQYLAGLSTLPGNPA,	
GEGAVQWMNRLIAFASRGNHV,	GSSYGFQYSPGQORVE,	ASCLRKLGPPLRVW,
and LTCGFADLMGY.		

29. The composition of claim 28, further comprising at least one epitope selected from the group consisting of LTDPSHITA, LADGGCSGAY, RMILMTHFF, VMGSSYGF, FWAKHMWNFI, LLFNILGGWV, IPFYGKAI, and VGIYLLPNR.

1 / 2

Monte Carlo population coverage analysis for HCV candidate epitopes



Plot of total frequency of genotypes as a function of the number of HCV candidate epitopes bound by HLA-A and B alleles, in an average population. Genotype values were derived by averaging the gene frequencies in Caucasian, North American, Black, Japanese, Chinese, and Hispanic populations. Also shown is the cumulative frequency of genotypes.

Using currently available HLA typing data, a residual fraction (about 15%) of the genes, in an average population, are unspecified. To arrive at 100% accounting of genes, a fraction of the residual has been added for each hit population cluster in proportion to the relative frequency of the cluster within the HLA specified population. One peptide, 24.0086, was not incorporated into the present analysis.

FIG. 1

2 / 2

HVC Minigene

CTL Epitopes

Kozak	SigSeq	Core 43	NS4 1590	NS3 1128	NS5 2611	Core 169	NS1/E2 632	NS4 1765	NS4 1863	Core 132
		1073.11	1013.02	1069.62	1090.02	1145.12	1073.13	24.0092	1073.10	1013.10
		A3	A2	A1	A2	B7	A3	A24	A3	A2

NS3.1252	NS4.1921	1437	NS5.2641	1466
1283.21	1283.44	35.0106	1283.55	35.0107
DR	DR	DR3	DR	DR3

HTL Epitopes

FIG. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/19774

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : A61K 38/00, 38/04, 38/08, 38/10, 39/29, 39/295

US CL : 514/2, 12, 13, 14, 15, 885; 424/185.1, 189.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/2, 12, 13, 14, 15, 885; 424/185.1, 189.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

MEDLINE, BIOSIS, EMBASE, DERWENT WPI, WEST 2.0, search terms: author names, hcv, pepid?, HLA, htl, cti,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim N
Y	WENTWORTH et al. Differences and similarities in the A2.1-restricted cytotoxic T cell repertoire in humans and human leukocyte antigen-transgenic mice. Eur. J. Immunol. 1996. Vol 26. pages 97-101, see entire document.	1-29
Y	US 5,736,142 A (SETTE et al.) 07 April 1998, see entire document.	1-29

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

B earlier document published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

**

later document published after the international filing date or prior date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X

document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y

document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document combined with one or more other such documents, such combinations being obvious to a person skilled in the art

*A**

document member of the same patent family

Date of the actual completion of the international search

19 SEPTEMBER 2000

Date of mailing of the international search report

17 OCT 2000

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks

Box PCT

Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

RON SCHWADRON

Telephone No. (703) 308-0196